NOV 19 1923

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

L ROOM LIBRARY E MICH.

BOARD OF EDITORS:

HAROLD ORDWAY RUGG, *Chairman.*
Lincoln School of Teachers College.
Teachers College, Columbia University.

JAMES CARLETON BELL,
Brooklyn Training School for Teachers.

FRANK NUGENT FREEMAN,
University of Chicago.

ARTHUR IRVING GATES,
Teachers College, Columbia University.

VIVIAN ALLEN CHARLES HENMON,
University of Wisconsin.

RUDOLF PINTNER,
Teachers College, Columbia University.

BEARDSLEY RUML,
Carnegie Corporation, New York City.

LEWIS MADISON TERMAN,
Leland Stanford University.

EDWARD LEE THORNDIKE,
Teachers College, Columbia University.

LAURA ZIRBES, *Assistant Editor.*
Lincoln School, of Teachers College.

VOL. XIV. NO. 8. NOVEMBER, 1923 $4.00 a year. 60c. per copy.

CONTENTS

AN EXPERIMENTAL STUDY OF THE GROWTH OF SOCIAL PERCEPTION. *Georgina Strickland Gates* 449

PERMANENCE OF HIGH SCHOOL LEARNING. *D. H. Eikenberry* . . 463

A STANDARDIZATION OF CERTAIN OPPOSITES FOR CHILDREN OF GRADE SCHOOL AGE. *Grace Arthur* 483

FURTHER NOTE UPON THE PROBABLE ERROR OF THE MEAN. *Truman L. Kelley* 497

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES. 499

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 503

Published Monthly Except June to August by

WARWICK and YORK, Inc.,

York, Pa. *Baltimore, Md.*

Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania, under the Act of March 3, 1879.

To be Published in November

A Laboratory Manual in the Psychology of Learning.

By William Henry Pyle.

Price $1.50; postage 8c.

Since the public school has as its purpose the direction and guidance of the learning of children, teachers should know as much as possible about the nature of the learning process and the laws of learning. If it is worth while to have laboratory courses in physics, chemistry, and the biological sciences, it must also be worth while to have laboratory courses in the psychology of learning. Sciences dealing with human nature have been too vague and general both as to the facts and their application. They must get closer to the facts; they must rely upon the laboratory. No amount of book study about human nature can take the place of even a few carefully devised and executed experiments. *A Laboratory Manual in the Psychology of Learning* is published in the hope of furthering and facilitating experiments in educational psychology. As now presented it represents the results of fourteen years of experimentation.

CONTENTS: Experimentation, The Mathematical Treatment of Data, Motor Learning—Trial and Error Type, Motor Learning, Semi-motor Learning, Retention of Motor Learning, A Study of Inhibition, Tachistoscopic Learning, Serial Learning, Associative Learning, Verbatim Learning of Poetry, Ideational Learning, Comparative Study of All the Experiments.

Warwick and York, Publishers, Baltimore

McConnell Made Maps

Teach History with good maps

After 30 years of experience in the manufacture of school maps we have scrapped our old maps and have produced the best maps made for the teaching of History.

The work on these maps has been done by the following authorities:

R. M. Tryon, Ph. D., University of Chicago
J. A. James, Ph. D., Northwestern University
A. G. Terry, Ph. D., Northwestern University
I. J. Cox, Ph. D., Northwestern University
C. L. Grose, Ph. D., Northwestern University
E. Lauer, Ph. D., Northwestern University
C. R. Fish, Ph. D., University of Wisconsin

Set 4 and Set 24—Ancient History
Set 5 and Set 25—Medieval and Modern History
Set 26 —Early European History
Other sets in preparation
Set 27 —Modern European and World History
Set 23 —American History

Set #3—Geography and United States History
Price, with adjustable steel stand, $38.00

Write for list of maps in these sets.

SEND IMMEDIATELY FOR WHAT YOU NEED. WE PAY THE TRANSPORTATION CHARGES. IF THE MAPS ARE NOT SATISFACTORY THEY MAY BE RETURNED AT OUR EXPENSE.

McCONNELL SCHOOL MAP CO.

213 Institute Place, Chicago, Ill.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XIV November, 1923 Number 8

AN EXPERIMENTAL STUDY OF THE GROWTH OF SOCIAL PERCEPTION[1]

GEORGINA STRICKLAND GATES

Barnard College, Columbia University

Ability to meet successfully social situations (social intelligence as it is sometimes called) depends, in part, on the capacity to make the appropriate tactful, courteous, aggressive, etc., reaction and, in part, on the ability to perceive accurately the conditions which are encountered. Of the latter requirement, adequate interpretation of the facial expressions of others forms an important ingredient.

That this ability increases with age is obvious. A very young child may mistake friendly behavior for threatening, or misinterpret fear as anger, or fail entirely to note signs of nervousness or excitement in those around him. An adult may be able to detect very slight evidence of chagrin or distaste or amazement.

It is the purpose of this paper to attempt a study of the genesis of the ability to interpret the emotional expressions of others. Photographs were used as the material because of their availability and of the possibility of exact standardization here. Six pictures were chosen from a set published by Ruckmick,[2] the basic of selection being the diversity of emotions represented and the relative ease of interpretation by adults, as determined by the performance of the individuals in Ruckmick's experiment, and by the judgment of 10 individuals in an advanced psychology class tested by the present author. Picture A (Plate I) shows a laughing face expressing joy, mirth, amusement; Picture B is posed for pain; Picture C expresses anger or defiance

[1] The experimental work of this study was performed by Helen Goldstone, Margaret Miller and Mary E. Langton. Miss Goldstone is responsible for the testing of Group II, Miss Miller of Group I, and Miss Langton of Group III.

[2] Ruckmick, C. A.: *Psychological Monographs* No. 136, 1921, pp. 30–35.

Picture D shows fear or horror; Picture E is scorn, contempt or disdain; and Picture F represents surprise, wonder or amazement.

The pictures were shown singly and individually to 458 children ranging in ages from 3 to 14, and varying greatly in social status. Three groups of children were tested; Group I was composed of 79 individuals from 3 to 6 years old, from the Kindergarten of a select private school; Group II of 265 children from the Kindergarten, first, second, third and fourth grades of a public school in a poor neighborhood of New York City, of which Irish, Jewish, and Italian children formed the majority; and Group III was made up of 30 children from a Settlement House in New York City, and 84 from the fifth and sixth grades of a public school which children of average social status attended. Table I shows the distribution of ages.

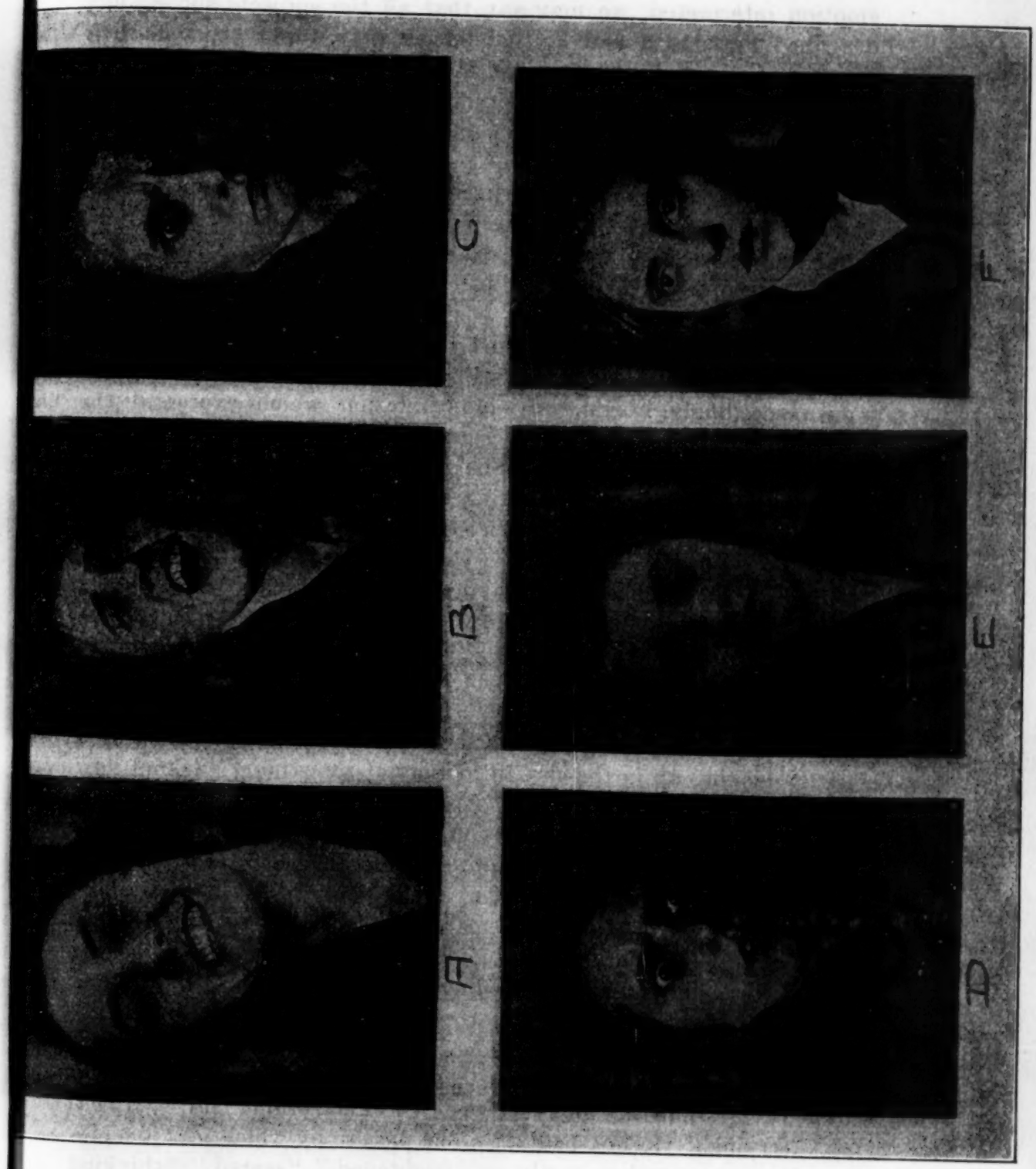
The child was asked in the case of each picture, "What is this lady doing?" and if this brought no response, or an ambiguous one, "What is the lady thinking about?" or "How does she feel?" or the subject was simply urged to tell more about it. Not mere understanding of the emotion then, but ability to convey the interpretation in some way to others was required. The response was in each case, of course, recorded verbatim.

The pictures were also presented to the members of an introductory class in psychology, in an attempt to obtain an adult standard of comparison. The purpose of the experiment was explained to these subjects, and they were asked to write down, in each case, the emotion or feeling represented.

TABLE I.—NUMBER OF CHILDREN TESTED IN VARIOUS AGE GROUPS

Age at last birthday	3	4	5	6	7	8	9	10	11	12	13	14
Group I.....	10	26	32	11								
Group II.....		14	53	47	50	52	31	7	4	4	3	
Group III.....							8	21	40	23	14	8
Total.....	10	40	85	58	50	52	39	28	44	27	17	8

Table I gives the results of the adult interpretations. A very close agreement is evident. If we score the responses somewhat liberally, disregarding differences in vocabulary, in understanding of the instructions, and variations in degree or in minor accompaniments of the


C
F
B
E
A
D

emotion interpreted, we may say that all the subjects successfully described Pictures A and E, that 98 per cent were correct in their interpretations of D, 95 per cent in C, 89 per cent in B and 84 per cent in F. Evidently these observers agree most closely on laughter and scorn, next on fear, next on anger, and least well on pain and surprise.

Lack of space prevents the printing of the responses made by each child to each picture. The interpretations made by Group I (the Kindergarten class of a private school) to Pictures A, C and F are, however, recorded in Table I to illustrate the kind of response obtained from the younger children.

Difficulties much more marked than in the case of adults occur when an attempt is made to evaluate the replies. Our standards were exceedingly liberal, any interpretation which expressed the general trend of feeling in the picture was counted as correct. In the first picture, such responses as "She's laughing," "happy," "smiling," "grinning," "looking at something nice," "thinking about her cute little baby," were obviously correct, and mere descriptions of the features ("showing her teeth," "cleaning her teeth," "putting her teeth in") such replies as "making faces" and the occasional "I don't know" or refusals to answer were, of course, unsuccessful responses. In the second picture, "cying," "hollering," "having bad dreams," "tooth-ache," "sees something she doesn't like," "mad," "sick," "sad," "pleading," "Ow," "she's cross," "she's frightened," were considered good interpretations, and "sneezing," "making a face," "looking up in the sky," "looking at something," "shouting," "showing her teeth," "singing," "staring," were expressions which did not seem to describe the emotion represented.

Pictures C, D, E and F, required a more specific interpretation, not merely pleasant or unpleasant feeling tone must be described, but a suggestion of the anger, fear, scorn, or surprise and wonder represented. For Picture C we counted as correct, "mad," "looking mean," "cross," "angry," "scolding," "looks furious," "like killing some one," "fighting," and as incorrect such interpretations as "laughing," "smiling," "scared," "singing," "sad," "coughing," "amazed," "crying," "making a funny face," "pouting," and even such partly correct expressions as "looks serious," "sees something bad," "frowning," "crying," "feeling badly." The picture by which fear is represented was reported correctly as "frightened," "scared," "thinking about a wild animal," "looks as if she saw something terrible,"

"horrified," and incorrectly as "laughing," "looking sad," "cross," "mad," "playing the piano," "scolding," "opening her mouth," "making a face," "sore," "surprised," "sick," "crying," etc. Scorn or contempt which was much more difficult for the children, yielded some replies which we might consider correct, as, "don't like anybody," "doesn't think much of a person," "don't care," "sees something that isn't nice," "disagreeable proud," "scornful," "disgusted," "jealous," "she's a tough one," "knows it all," "shows off," and obviously inadequate expressions as "dressing," "looking at something," "looking nice," "silly," "laughing," "likes somebody," "looks aside," "sorry," "happy," "miserable," "pensive" "mad," "crying." "Surprise," "wondering," "astonished," "staring," "something suddenly came," "saw something," were correct interpretations of the last picture, and "washing teeth," "making a face," "scolding," "hollering," "praying," "breathing," "smiling," "looking wise," "looking stern," "sick," "posing," "just over a fit," "not well," "angry," and replies which denoted fear alone without the element of surprise, were among the incorrect responses.

It is very easy to quarrel with these classifications or to question the individual interpretation of replies made by the examiners. In the case of Picture B, for example, though the original intention was to limit the acceptable responses to pain, sadness, grief, crying, and the like, so many of the older children interpreted this as anger or fear, that it was decided to add these to the list of successful responses. For the picture of scorn, the reply "jealous" frequently given, was after some debate, admitted as a correct reaction. A large number of cases occur where the proper scoring of individual responses was somewhat doubtful.

Taking the scores then, for what they are worth, and realizing the difficulties of interpretation which the questioner often met, we find, according to Tables IV and V, a gradual increase in ability to interpret each picture as we pass from the youngest to the oldest children. Occasional variations occur which we may and will explain, as due to differences in number or ability of the age groups. The general trend is upward. Laughter is understood by more than half of the children whose age at the last birthday was three; pain is correctly interpreted by more than half of those from ages 6 years 0 months to 6 years 11 months; anger is understood at 7; fear at 10; surprise at 11; and scorn is described by only 43 per cent of the 11-year-old children, though all the adults in the superior group tested, understood this picture.

TABLE II.—INTERPRETATIONS OF THE PICTURES MADE BY 36 COLLEGE STUDENTS

Picture A		Picture B		Picture C		Picture D		Picture E		Picture F	
Interpretation	No. of cases	Interpretation	No. of cases	Interpretation	No. of cases	Interpretation	No. of cases	Interpretation	No. of cases	Interpretation	No. of cases
Amusement	11	Amazement	1	Anger.....	28	Astonishment .	1	Contempt.....	15	Amazement....	3
Happiness.....	4	Anguish.....	6	Concentration..	2	Fear.....	19	Cynical expression.....	1	Astonishment .	4
Joy.....	5	Calling to someone Upstairs..	1	Disapproval....	1	Horror.....	12	Disdain.....	2	Fear.....	3
Laughter.....	6	.....	..	Hate.....	2	Terror.....	4	Disgust.....	6	Hope.....	1
Mirth.....	6	Grief.....	1	Mad.....	1			Scorn.....	12	Horror..	1
Pleasure.....	4	Horror.....	2	Repugnance....	1					Surprise.....	19
		Forced laughter	2	Stubbornness...	1					Wonder.....	4
		Misery.....	1							Vacant expression.....	1
		Pain.....	22								

When we group the children by grades (Table VI), we find, if we use again the 50 per cent mark as our measure, that more than half of the Kindergarten children understood the laughter picture, more than half of Grade I the pain picture, Grade III anger, Grade IV fear, Grade VI surprise or wonder, and the portrayal of scorn is understood by 43 per cent of Grade V.

The evidence from this testing shows, then, that the probable order of difficulty for the pictures (from least to greatest) is for children, laughter, pain, anger, fear, surprise, and scorn. This differs from that found for adults (Table II) where the order is laughter and scorn (100 per cent) fear, anger, pain, and surprise. The greatest discrepancy occurs in the case of scorn which is easy for the superior adult and extremely difficult for the child, and pain, which the child readily interprets and which the adults seem to feel to be inadequately represented.

It may be that differences occur which are due to the social status of the individual tested. We might expect that the children from the private school would do better because they are, in general, brighter, and because their vocabulary is presumably more extensive; or we might expect that the children from poorer neighborhoods, of foreign parentage would have more opportunity to observe expressions of emotions both at home and in the movies which they attend more frequently than their more wealthy competitors, and so might show greater ability along this line. An examination of Table IV shows that when we compare children of like ages in Groups I and II, that Group II is better in interpreting the pain picture and Group I slightly better in interpretations of the other five. The small size of the differences and the fact that different examiners questioned the two groups, should be taken into account when the statement is made that in general the children from the private school did better.

Sex differences, if they exist, should be described here as elsewhere. Perhaps the superior "social tact" attributed to women is due to better ability to interpret the emotional expressions of others. A tabulation according to sex and age shows no differences. At the ages of four, five, and nine, the girls are slightly superior; at five, six, seven and eight, the boys. As the differences are small as well as inconsistent, chance or other factors (as differences in intelligence) may be invoked to explain them. There appears to be no difference in the reactions of the two sexes to the various pictures except in the case of the representation of fear where the boys are always superior.

TABLE III.—INDIVIDUAL RESPONSES MADE TO PICTURE A (LAUGHTER) BY GROUP I

Response	Number of times given by children of each age			
	3 yr.	4 yr.	5 yr.	6 yr.
Laughing	6	18	23	5
Smiling	1	2	1	1
Putting teeth together	2			
Putting teeth in	1			
Rough	..	1		
Showing teeth	..	3		
No response	..	1	1	
Making faces	..	1		
Grinning	..	..	1	
Looking at something	..	..	1	
Tipping over head	..	..	1	
Looking at her teeth	..	..	1	
Thinking of happy things	..	..	1	
Thinking of something nice	..	..	1	
Sees doggie	..	..	1	
Having her picture taken	..	..	..	1
Cleaning her teeth	..	..	..	1
Happy	..	..	..	1
Sees something she likes	..	..	..	1
Thinking about her cute little baby	..	..	..	1

TABLE III.—(*Continued*)

Response	Number of times given by children of each age			
	3 yr.	4 yr.	5 yr.	6 yr.
Mad	2	4	7	2
Looking mean	1			
Not laughing	1			
Opening her eyes	1	1	1	
Frowning	1	..	4	
"Eyes"	..	1		
Crying	..	1	..	1
Cross	..	9	3	1
Making a funny face	..	5	..	1
Nothing	..	1		
Sitting	..	1		
Walking	..	1		
Feeling badly	..	..	1	
Making a face	..	..	5	
Looking at a picture	..	..	1	
Pouting	..	..	1	
Scolding	..	..	1	1
Looks serious	..	..	1	
Looks furious	..	..	1	
Looking	..	..	1	
Sees me	..	..	1	
Angry	..	..	1	
Sees bird	..	..	2	
Angry naughty lady	..	..	..	1
Not happy	..	..	..	1
Sees something bad	..	..	..	1
Sees rat	..	..	..	1
Thinking about something mad	..	..	..	1
"I don't know"	4	2	1	1

TABLE III.—(*Continued*)

Response	Number of times given by children of each age			
	3 yr.	4 yr.	5 yr.	6 yr.
Staring	1			
Looks mean	1			
Mouth open	2	1	4	
Not laughing	1			
Nothing	1	2	1	
Something funny	1			
Looking at something	..	2	4	
Just looking	..	7	2	2
Washing teeth	..	1		
Looking sad	..	1		
Being friendly	..	1		
Looks wild	..	1		
Sitting	..	1		
Walking	..	1		
Isn't making a face	..	1		
Talking	..	1	1	2
Cross	..	1	1	
Making a face	..	1	2	
Just all alone	..	..	1	
Dogs trying to get in	..	..	1	
Sees animals	..	..	2	2
Scolding	..	..	1	1
Looks serious	..	..	1	
Showing teeth	..	..	1	
Smiling	..	..	1	
Mad	..	..	1	
Round eyes	..	..	1	
Opening mouth and eyes	..	..	1	
Frightened	..	..	1	
Hears something	..	..	1	
Laughing	..	..	1	
Sees something she likes	..	..	..	1
Surprised	..	..	..	1
"I don't know"	3	4	3	1

TABLE IV.—PERCENTAGE OF SUCCESSFUL RESPONSES AT VARIOUS AGES, FOR CHILDREN IN THE THREE GROUPS

Age at last birthday.....	3	4	5	6	7	8	9	10	11	12	13	14	Adult
Picture A.....													
Group I.....	.70	.76	.84	.82									
Group II.....	...	.57	.83	.93	.94	.90	.93	100	100	100	100	...	.100
Group III.....	...	...	...	...	...	...	100	100	97	100	100	100	
Picture B.....													
Group I.....	.40	.40	.40	.45									
Group II.....	...	.43	.42	.62	.72	.43	.64	.72	100	.50	100	...	.89
Group III.....	...	...	...	...	...	...	.63	.62	.77	.92	.71	100	
Picture C.....													
Group I.....	.30	.50	.40	.54									
Group II.....	...	.21	.23	.37	.52	.61	.74	.85	.50	100	100	...	.95
Group III.....	...	...	...	...	...	...	.62	.80	.85	.92	.57	.87	
Picture D.													
Group I.....	.00	.03	.12	.36									
Group II.....	...	.07	.15	.13	.10	.32	.54	.43	.75	100	.33	...	.98
Group III.....	...	...	...	...	...	...	.12	.61	.77	.86	.64	.75	
Picture E.....													
Group I.....	.00	.00	.00	.09									
Group II.....	...	.00	.00	.00	.04	.08	.13	.29	.25	.00	.00	...	100
Group III.....	...	...	...	...	...	...	...	.48	.45	.52	.22	.25	
Picture F.....													
Group I.....	.10	.00	.00	.09									
Group II.....	...	.00	.00	.00	.00	.00	.07	.00	.00	.00	.00	.,.	.84
Group III.....	...	...	...	...	...	...	.00	.05	.62	.60	.50	.75	

TABLE V.—PERCENTAGE OF SUCCESSFUL RESPONSES FOR CHILDREN OF VARIOUS AGES, COMBINING THE THREE GROUPS

Number of cases	10	40	85	59	55	58	39	28	44	27	17	8	36
Ages	3	4	5	6	7	8	9	10	11	12	13	14	Adult
Picture A	.70	.70	.84	.89	.94	.90	.95	100	.98	100	100	100	100
Picture B	.40	.40	.44	.60	.72	.43	.64	.64	.79	.74	.77	100	89
Picture C	.30	.40	.29	.40	.52	.61	.72	.82	.82	.93	.65	.87	95
Picture D	.00	.05	.13	.17	.10	.32	.46	.57	.77	.89	.59	.75	98
Picture E	.00	.00	.00	.02	.04	.08	.10	.42	.43	.41	.18	.25	100
Picture F	.10	.00	.00	.02	.00	.00	.05	.04	.57	.52	.41	.75	84

The results of the experiment indicate in general, then, the feasibility of investigating the growth of social perception. They point to age differences in ability to interpret various emotional expressions, differences which we may attribute either to variations in truthfulness of the actual representations used, or to an actual change in capacity through which first laughter, then pain, then anger, fear, surprise, and scorn are recognized, understood, and made capable of description. That variations in difficulty of the pictures is not the entire cause of the differences is indicated by the results obtained from superior adults where the order of difficulty is not identical with that for children. Sex differences are not manifest, and differences caused by social status are, if they occur, slight.

It is possible that some use may be made of the age and grade norms of which these tables may be a starting point. We might say that inability to interpret correctly the pictures understood by more than half of the children of a given individual's age or grade, indicates a certain lack of social perception. Just what the significance of this incapacity is, cannot of course, be determined until the results of the test are compared with the data obtained from other measures. Correlations with social maturity (as judged by teachers) physical maturity, intelligence, etc. are now being prepared. Measurement by such a test of delinquents, and others whose social adjustment is obviously inadequate, has been suggested.

TABLE VI.—PERCENTAGE OF SUCCESSFUL RESPONSES GIVEN BY VARIOUS GRADES

Grade	Kindergarten (Public School)	Grade I (Group II)	Grade II (Group II)	Grade III (Group II)	Grade IV (Group II)	Grade V (Group III)	Grade VI (Group III)
Number cases	67	43	58	60	42	23	81
Picture A	77	93	95	83	100	100	100
Picture B	46	67	60	51	74	78	91
Picture C	22	35	45	61	83	78	81
Picture D	13	11	08	28	64	83	85
Picture E	00	00	02	08	17	43	41
Picture F	00	00	00	00	04	30	65

PERMANENCE OF HIGH SCHOOL LEARNING

D. H. EIKENBERRY

Graduate Student, Teachers College Columbia University

I. Introductory

The purpose of the present study was to determine as accurately as possible the permanence of learning in certain subjects studied in high school but not continued in higher institutions. The subjects of the study were the members of two senior classes in Educational Psychology, one in Rutgers College and the other in the New Jersey College for Women. The class in Rutgers was composed of 16 men and the class in the College for Women of 18 women.

Early in the fall semester a survey of the two classes was made to find out what subjects could be used in the study and the number of men and women available in each subject out of the total of 34 students. The available subjects and the number of students for each are given in Table I below:

Table I.—Subjects Studied in High School but not Continued in College

Subject	Semesters in H. S.	Men	Women	Total
Latin	2	0	1	1
	4	3	2	5
	6	3	1	4
	8	4	2	6
Total		10	6	16
Ancient History	1	0	1	1
	2	9	7	16
Total		9	8	17
United States History	1	0	2	2
	2	6	7	13
Total		6	9	15
Geometry	2	2	11	13
	3	3	2	5
Total		5	13	18
Physics	2	7	11	18
Chemistry	2	8	9	17

The complete plan for the study as finally formulated included the following:

1. Rough determination of achievement in high school by securing final semester marks in each subject.
2. Determination of present status in each subject by use of standard tests.
3. Rating of high school subjects as to interest manifested in them while in high school by use of a five point scale—A B C D E.
4. Measurement of the intelligence of the students.

With these data in hand it was possible to make the following specific studies:

1. Comparison of scores earned on the standard tests with the marks earned while in high school.
2. Comparison of the scores earned on the standard tests with norms for high school pupils.
3. Correlation of intelligence with test scores.
4. Correlation of intelligence with school marks.
5. Correlation of intelligence with interest in each subject.
6. Correlation of test scores with school marks.
7. Correlation of test scores with interest ratings.
8. Correlation of school marks with interest ratings.
9. Correlation of test scores with lapse of time.
10. Sex differences.
11. Comparison of subjects as to permanence of learning.

The final semester marks were secured from the principals of the high schools represented by the members of the two classes. A blank form for each student was sent to the principal who was asked to give the final semester mark, the dates the subjects were taken in high school, and the passing mark of the school at the time the marks were earned. These blanks were returned by all principals with all the data asked for. The passing mark in all schools was found to be 70.

In determining the present status of each student in the subjects not pursued since leaving high school it was decided to make use of the following tests: Henmon, Latin Test 1, Vocabulary and Sentences; Sackett, Ancient History Scale; Van Wagenen, American History, Information Scale A; Barr, Diagnostic Tests in American History; Minnick, Geometry Test A; Starch, Physics Test; Bell, Chemistry Test. All of these tests were devised for high school use except the Van Wagenen American History Scale and standard scores were found with which to compare the scores earned by the college students. The

Van Wagenen Test was given to high school seniors in two New Jersey high schools as explained later for the purpose of securing scores with which to compare. All of the tests were given during November and December. The students were given no warning beforehand so it was not possible for them to cram quickly on the subjects and thus earn a higher score than they otherwise would.

In order to secure the interest ratings the students were asked to fill out in class a blank listing the subjects of the study, so arranged that they could check their interest while in high school in columns headed A B C D E. The instructions for filling this blank were as follows:

"In the proper space below please check your interest, while in high school, in the respective subjects by checking in the proper column according to the following plan: If the subject was your favorite, check under column A; if not your favorite but above average in interest, check under B; if only of average interest, check under C; if of less than average interest but not the subject of least interest, check under D; if the subject was detested, check under E. These interest ratings were secured before any of the tests were given in order that they should not be influenced by experiences with the test. In order to get a check on the reliability of the ratings the students were asked to rate Chemistry at the time the Bell Chemistry Test was taken. No mention was made of the former rating which was made about six weeks earlier. Of the seventeen students taking the Chemistry Test and rerating for interest there was found to be exact agreement with the first rating in all cases except one and in that case the change was one letter only.

The intelligence of the students was measured by a battery of three standard tests: namely, Army Alpha, Test VI of the Bureau of Personnel Research commonly known as "Scrambled Alpha," and the Terman Group Test of Mental Ability with time limits cut in two. The results of these three tests are given in Tables II, III, and IV. Table II shows the results for the sixteen men, giving the score earned on each test and the rank each man holds in his group. Table III gives the same for the eighteen women and Table IV combines the two sexes. In these tables and all following tables men are represented by single capital letters and women by double capitals.

That the battery of tests slightly favored the men is indicated by the average scores in Tables II and III. For purposes of comparison these are given below:

	Army Alpha	Test VI	Terman	Average
Men	172.0	156	167.6	166
Women	169.8	150	168.0	161
Both	171.0	153	167.8	163

Table II.—Intelligence Scores, Sixteen Men

Pupil	Army Alpha		Scrambled Alpha		Terman		Average	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank
A	176	7½	157	9½	164	11	166	8½
B	188	3	175	1	182	3½	182	2
C	176	7½	157	9½	164	11	166	8½
D	158	13	163	5	168	9	163	11½
E	193	1	174	2	182	3½	183	1
F	156	14	127	16	133	16	139	16
G	168	10½	130	15	172	7	163	11½
H	168	10½	167	4	150	14	163	11½
I	179	6	171	3	186	2	179	3
J	181	4½	156	11	187	1	175	5
K	166	12	158	8	164	11	163	11½
L	181	4½	161	7	171	8	171	6
M	192	2	162	6	178	6	177	4
N	175	9	152	12	180	5	169	7
O	155	15	143	14	154	13	151	14
P	141	16	144	13	147	15	144	15
Average	172		156		167.6		166	

Correlation of Scores (Spearman's Footrule Method)

	r
Army Alpha and Scrambled Alpha	.554
Army Alpha and Terman	.711
Army Alpha and Average	.926
Scrambled Alpha and Terman	.458
Scrambled Alpha and Average	.677
Terman and Average	.818

This comparison shows that the women are three points behind the men in Army Alpha, six points behind in Test VI and one point ahead in the Terman Test. Judging from the results of the Terman Test it would seem that the women are equal in ability to the men. In all likelihood the somewhat poorer results on Army Alpha and Test VI are due to the fact that these tests were devised primarily for men.

No attempt was made to reduce the intelligence scores to mental ages or intelligence quotients as they are used only in ranking the students as to mental ability.

Coefficients of correlation were found by the Spearman Footrule Method of each test with the two others and with the average of the three. These correlations are given below, Tables II, III and IV, and for purposes of comparison are brought together on page 469.

TABLE III.—INTELLIGENCE SCORES, EIGHTEEN WOMEN

Pupil	Army Alpha		Scrambled Alpha		Terman		Average	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank
AA	169	8	164	3	179	3	171	4½
BB	189	1	152	8½	171	8	171	4½
CC	162	13½	145	14½	171	8	159	12
DD	159	16½	119	18	137	17	138	18
EE	166	9	135	16	161	12	154	14
FF	164	11½	159	6	171	8	165	8½
GG	161	15	145	14½	132	18	146	16
HH	159	16½	150	10	172	6	160	11
II	164	11½	169	2	151	13	161	10
JJ	187	2	153	7	183	2	174	2
KK	162	13½	148	12	141	15	150	15
LL	179	5	171	1	177	4	176	1
MM	165	10	146	13	185	1	165	8½
NN	179	5	160	5	169	11	169	6
OO	180	3	163	4	174	5	172	3
PP	179	5	149	11	170	10	166	7
QQ	156	18	126	17	143	14	142	17
RR	176	7	152	8½	138	16	155	13
Average.....	169.8		150		168		161	

CORRELATION OF SCORES (SPEARMAN'S FOOTRULE METHOD)

	r
Army Alpha and Scrambled Alpha.....	.592
Army Alpha and Terman.....	.472
Army Alpha and Average.....	.836
Scrambled Alpha and Terman.....	.486
Scrambled Alpha and Average.....	.810
Terman and Average.....	.742

COMPARISON OF CORRELATIONS

	MEN	WOMEN	BOTH SEXES
Army Alpha and Scrambled Alpha	.554	.592	.618
Army Alpha and Terman	.711	.472	.593
Army Alpha and Average	.926	.836	.896
Scrambled Alpha and Terman	.458	.486	.500
Scrambled Alpha and Average	.677	.810	.791
Terman and Average	.818	.742	.844

TABLE IV.—INTELLIGENCE SCORES, MEN AND WOMEN

Pupil	Army Alpha		Scrambled Alpha		Terman		Average	
	Score	Rank	Score	Rank	Score	Rank	Score	Rank
A	176	14	157	15½	164	22	166	15
B	188	4	175	1	182	5½	182	2
C	176	14	157	15½	164	22	166	15
D	158	30	163	8½	168	20	163	20½
E	193	1	174	2	182	5½	183	1
F	156	31½	127	32	133	33	139	33
G	168	18½	130	31	172	12½	163	20½
H	168	18½	167	6	150	27	163	20½
I	179	10½	171	3½	186	2	179	3
J	181	6½	156	17	187	1	175	6
K	166	20½	158	14	164	22	163	20½
L	181	6½	161	11	171	15½	171	10
M	192	2	162	10	178	9	177	4
N	175	16	152	20	180	7	169	12½
O	155	33	143	29	154	25	151	28
P	141	34	144	28	147	28	144	31
AA	169	17	164	7	179	8	171	10
BB	189	3	152	20	171	15½	171	10
CC	162	25½	145	26½	171	15½	159	25
DD	159	28½	119	34	137	32	138	34
EE	166	20½	135	30	161	24	154	27
FF	164	23½	159	13	171	15½	165	17½
GG	161	27	145	26½	132	34	146	30
HH	159	28½	150	22	172	12½	160	24
II	164	23½	169	5	151	26	161	23
JJ	187	5	153	18	183	4	174	7
KK	162	25½	148	24	141	30	150	29
LL	179	10½	171	3½	177	10	176	5
MM	165	22	146	25	185	3	165	17½
NN	179	10½	160	12	169	19	169	12½
OO	180	8	163	8½	174	11	172	8
PP	179	10½	149	23	170	18	166	15
QQ	156	31½	126	33	143	29	142	32
RR	176	14	152	20	138	31	155	26
Average	171		153		167.8		163	

CORRELATION OF SCORES (SPEARMAN'S FOOTRULE METHOD)

	r
Army Alpha and Scrambled Alpha	.618
Army Alpha and Terman	.593
Army Alpha and Average	.896
Scrambled Alpha and Terman	.500
Scrambled Alpha and Average	.791
Terman and Average	.844

II. PERMANENCE OF HIGH SCHOOL LEARNING AS DETERMINED BY COMPARISON OF SCORES EARNED ON STANDARD TESTS WITH HIGH SCHOOL MARKS AND NORMS FOR HIGH SCHOOL PUPILS ON STANDARD TESTS

The first difficulty encountered in a study of this kind is that of determining with any degree of accuracy the amount of any subject ever known by the subjects of the study. Scientific accuracy would require that initial tests must have been given to all the students at the time they had completed the study of each subject in high school. Lacking these initial tests the only thing we can do is to compare the results on the tests given with high school marks and with standard scores for high school pupils where such exist. Using McCall's formula for a one group experimental problem:

$$S \ . \ . \ . \ IT—EF—FT—C$$

Where S is the experimental group, IT the initial test, EF the experimental factor, FT the final test and C the change we have EF, the experimental factor, namely, the lapse of time since the studies were completed in high school and FT, the final tests. It is necessary for us then to assume from high school marks and standard scores what the IT in each case would have been had it been given to the students while in high school.

The second difficulty in the study is that of determining the effect of incidental learning. In the case of each high school subject care was taken not to include any students who had continued the study in college. But this does not guarantee that the subjects learned in high school have not been kept up more or less by incidental learning in other subjects and in ordinary ways out of classes. In the above formula C, the change between initial and final test is undoubtedly smaller on account of incidental learning than it would be if we could eliminate this factor. This seems to be especially true in the case of United States History as will be pointed out later.

TABLE V.—HIGH SCHOOL MARKS AND SCORES ON STANDARD TESTS

Pupil	Latin		Ancient history		U. S. history			Geometry		Physics		Chemistry	
	H. S.	Test	H. S.	Test	H. S.	Test I	Test II	H. S.	Test	H. S.	Test	H. S.	Test
A	90	66	87	205	..	....	..	..	....	..	....	88	55
B	..	..	....	...	..	....	..	88	78				
C	78	89	....	...	..	....	..	79	68				
D	..	..	....	...	..	....	..	..	....	90	29	87	26
E	90	80	....	...	90	79	80						
F	75	34	74	79	..	....	..	..	....	85	0	90	10
G	80	44	85	234	75	72	57	70	63				
H	..	..	82	118	..	....	..	..	....	78	15	74	30
I	74	48	....	...	..	....	..	..	....	73	10	71	24
J	..	..	94	241	93	75	58						
K	..	..	....	...	75	46	39	..	....	65	4		
L	73	42	80	159	85	66	58	84	80				
M	81	85	....	...	..	....	..	..	....	..	....	91	44
N	74	62	74	80	76	64	63	75	80				
O	..	..	....	...	..	....	..	..	....	..	....	75	40
P	76	25	92	190	..	....	..	..	....	80	0		
Average.....	79	58	83½	163	82	67	59	79	74	79	9.7	82	33
AA	..	..	....	...	83	51	52	..	....	95	8		
BB	..	..	....	...	..	....	..	..	....	80	39		
CC	..	..	80	134	74	18	53	89	29				
DD	..	..	95	167	95	40	47	..	....	..	....	95	12
EE	89	45	....	...	..	....	..	73	49	74	8	73	14
FF	73	76	....	...	..	....	..	75	51				
GG	89	64	90	139	..	....	..	80	60				
HH	..	..	....	...	91	45	53	96	59	87	11	90	16
II	..	..	84	92	..	....	..	79	25	..	....	73	10
JJ	..	..	....	...	..	....	..	..	....	93	24		
KK	..	..	93	120	95	43	56	96	38	95	7		
LL	92	63	....	...	..	....	..	95	37				
MM	92	38	88	100	87	28	55	..	....	81	18		
NN	..	..	92	187	..	....	..	96	72	84	10	91	43
OO	..	..	....	...	85	51	57	77	24	73	4	73	24
PP	78	16	....	...	82	53	51	85	36	..	....	70	16
QQ	..	..	90	115	83	35	52	83	38	..	....	78	25
RR	77	41	....	...	..	....	..	80	63	78	8	79	20
Average.....	84	49	89	132	86	40½	53	84	44.7	84	13.7	80	20
Standard.....	..	100	....	171	..	55½	43	..	62.5	..	50	..	59

Table V above shows the marks earned in high school and the scores on the standard tests. The upper half of the table (letters A to P) includes the sixteen men; the lower half (letters AA to RR) includes the eighteen women. Average high school marks and test scores are given for each sex. The Latin scores are not the raw scores earned but revised scores as explained on page 471. Scores under Test I of United States History are on the Van Wagenen Scale and scores under Test II are on the Barr Test.

Table V above gives the marks earned in high school in the six subjects of the study and the scores earned in the seven standard

tests used. If high school marks were highly reliable and meant the same thing in one subject as in another, and in one school as in another, it would be possible to take the distribution of marks in each subject and transmute into scores on the standard tests. It would then be possible to compare the transmuted marks with the scores earned on the standard tests and we would be able to compute the percentage of subject matter retained provided the standard tests were scaled from a reliable zero point. Correlations of high school marks were so low that no attempt was made to transmute the marks. The coefficients are given below to show why the high school marks could not be transmuted into scores:

	r
Latin marks and scores on Henmon Test	.192
Ancient History marks and scores on Sackett Test	.541
U. S. History and scores on Van Wagenen Test	.018
U. S. History and scores on Barr Test	.192
Geometry marks and scores on Minnick Test	.176
Physics marks and scores on Starch Test	.275
Chemistry marks and scores on Bell Test	.259

(Spearman's Footrule method used in computing r's)

The best comparisons we are able to make for purposes of determining retention are between the scores earned on the standard tests and the standard scores for high school pupils in the same tests. It is unfair, however, to make the comparisons with the median scores of high school pupils as college seniors are a more highly selected group. In order to allow for selection we will probably not be far wrong if we make our comparisons with the upper half only of high school pupils. Our problem is then to compare the scores of college seniors with the average or median of the upper half of high school pupils. From the published results of the tests used in the study it was possible to get the 75 percentile only in the case of the Minnick Geometry Test. Results of the Bell Chemistry Test in fourteen Texas high schools give the 70 and 80 percentiles from which the 75 percentile may be estimated accurately enough for our purposes. The Van Wagenen Test was given to 97 pupils in the New Brunswick and Elizabeth high schools so it was possible to get the average ability of the upper half of the group. Table VI gives the standard score for each test used, the 75 percentile where it is known, the average scores of the college seniors and the median number of years since completing the study of the subject.

TABLE VI.—COMPARISON OF RESULTS ON STANDARD TESTS WITH STANDARD SCORES FOR HIGH SCHOOL PUPILS AND WITH MEDIAN SCORES OF THE UPPER HALF

Subject	Standard score	75 Percentile	College seniors			Years ago
			Men	Women	Both	
Latin	100.0	....	58.0	49.0	54.0	3½
Ancient History	171.0	....	163.0	132.0	147.5	5½
U. S. History (Van W.)	55.5	62.0	67.0	40.5	51.0	3½
U. S. History (Barr)	43.0	....	59.0	53.0	55.0	3½
Geometry	62.5	72.9	74.0	44.7	53.0	4½
Physics	50.0	....	9.7	13.7	12.2	4½
Chemistry	49.0	67.0	33.0	20.0	25.5	4½

Figure 1 presents the same facts graphically. In all cases the standard score is assumed to be 100 and the score that percentage of 100 which the score in Table VI is of the standard in Table VI.

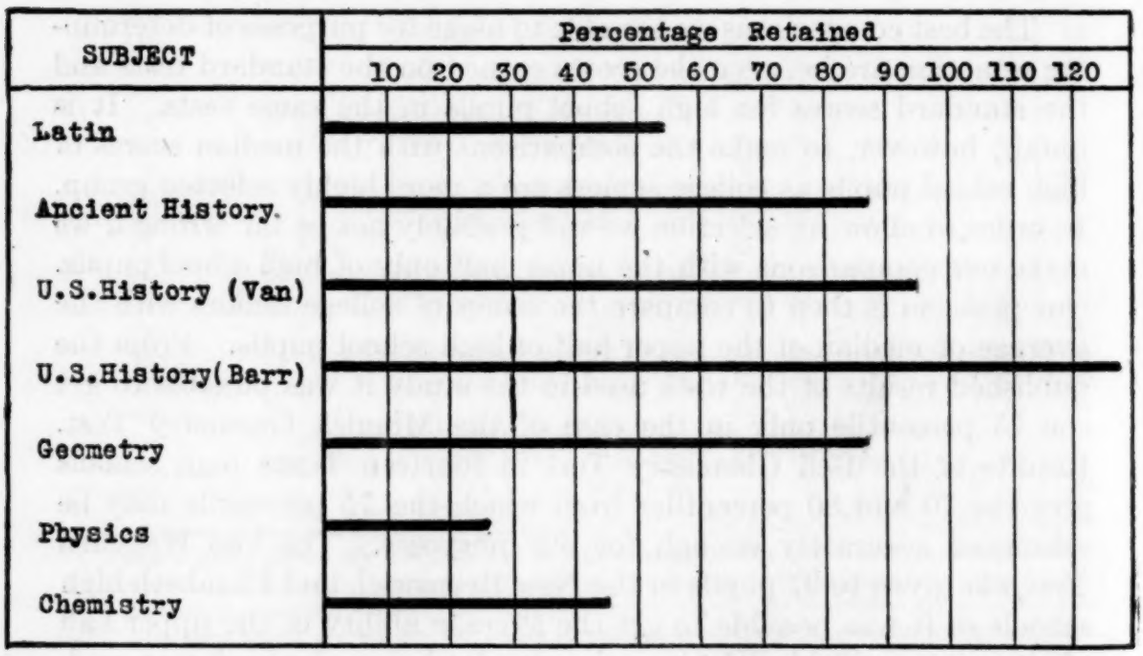


FIG. 1.

From Table VI and Fig. 1 it is seen that retention of high school learning seems to be highest in United States History, Ancient History,

and Geometry, and lowest in Physics, Chemistry, and Latin. As judged from test results it would seem that college seniors who have not studied United States History for three or four years possess almost as much as the average high school pupils retain at the close of their study of the subject. On the Van Wagenen Test the median for 97 high school seniors who were just completing their study of United States History was 55, and the 75 percentile, 62. The average score of the men on this test was 67, of the women 40½, and for both sexes 51. If we assume that the ability of the college seniors, while in high school, is represented by the 75 percentile then our test results show the men to be five points above the 75 percentile and the women 21½ points below. Both sexes combined are 11 points below. On the Barr Test the standard is 43. The average of the men is 59, of the women 53, of both sexes 55. The 75 percentile is unknown, but it is probably not far if any above the scores of the men and women.

The high retention found in the case of United States History is no doubt due very largely to the factor of incidental learning. Both in and out of school college students have numerous opportunities to keep the main facts of American History in mind. Courses in Economics, Sociology, American Literature, and the like in school, and newspaper and magazine reading outside of school, to say nothing of many other contributing factors, serve to keep the bonds strong enough for recall. The same thing is true but to a lesser degree in Ancient History. College courses in Greek and Latin Literature, Art, Philosophy, and the like contribute much in keeping the chief facts from being forgotten. The comparatively high retention in Geometry may be due partly to incidental learning in other mathematics courses but is perhaps largely due to the fact that not memory alone was needed to solve the problems of the test but reasoning also. The Minnick Geometry Test A which was used in this study consists of five problems in construction. In each case the proposition is stated and the pupil is required to make the drawing. It is not a case of sheer memory, but of ability to reason out the answer. This same condition holds also in the case of the Barr Diagnostic Test in American History. It is so constructed as not to test information alone but to test also one's ability to reason with the facts of history. On this test as Fig. 1 shows the college seniors far exceeded the norm for high school pupils.

As judged from test results college seniors retain about half of their high school Latin. But here again it is likely that incidental learning is an important factor. College studies in English and Romance Languages may serve to keep the essentials of Latin in mind.

The much lower retention in the case of Physics and Chemistry is probably due to the fact that the students have had few if any opportunities to apply the knowledge or to review it since leaving high school. A great deal of science instruction is formal in which mastery of the text and performance of a required number of cut and dried experiments constitute the sole requirements. Pupils too often do not learn the relationship between the facts of science and the world about them and as a consequence the larger part of the facts once known soon fade away.

A detailed analysis was made of the right and wrong responses on the Henmon Test, Van Wagenen Test, Bell Test, Starch Test and the Sackett Test. Space does not permit giving these tabulations in full, but in Table VII are given representative samples from the five tests.

TABLE VII.—NUMBER OF CORRECT RESPONSES ON REPRESENTATIVE QUESTIONS IN FIVE STANDARD TESTS

TEST AND NO. OF QUESTION	STATEMENT OF THE QUESTION, LITERALLY OR SUBSTANTIALLY	TOTAL IN GROUP	TOTAL RIGHT
Latin			
Henmon, Vocab.			
5	Dico	17	17
10	Non	17	11
15	Spes	17	9
20	Virtus	17	10
25	Vita	17	14
30	Adventus	17	4
35	Appello	17	16
40	Nam	17	6
45	Tum	17	5
50	Quisque	17	1
Henmon, Sentence			
5	His rebus cognitis, totus exercitus prima luce discesserunt	17	4
10	Locus castris muniendis facilis nobis deligendus est	17	1

TABLE VII.—NUMBER OF CORRECT RESPONSES ON REPRESENTATIVE QUESTIONS IN FIVE STANDARD TESTS—(*Continued*)

TEST AND NO. OF QUESTION	STATEMENT OF THE QUESTION, LITERALLY OR SUBSTANTIALLY	TOTAL IN GROUP	TOTAL RIGHT
AMERICAN HISTORY			
Van Wagenen			
5	By what people was our Thanksgiving Day custom started?	15	6
12	What was Henry Hudson looking for when he sailed up the Hudson River?	15	6
17	What group of Indian Tribes lived in Western New York?	15	9
27*d*	Who was president of the Southern Confederacy?	15	7
27*i*	Who secured adoption of the Missouri Compromise?	15	2
27*j*	Who was the first Chief Justice of the Supreme Court?	15	4
34*a*	Who was president during the Mexican War?	15	0
34*c*	Who was president when Florida was purchased?	15	2
ANCIENT HISTORY			
Sackett			
I 1	What was Hannibal noted for?	17	11
7	What was Attila noted for?	17	10
II 1	Name a man noted as an orator	17	16
4	Name a man noted as an historian	17	8
III 1	Historical significance of battle of Tours	17	5
4	Historical significance of Council of Nicaea	17	3
10	Historical significance of Peloponnesian War	17	2
V 1	Approximate date of fall of Rome	17	5
6	Approximate date of Hejira	17	0
10	Approximate date of establishment of Roman Empire	17	2
VI 1	Contribution of Greeks to civilization	17	17
3	Contribution of Phoenicians to civilization	17	5
5	Contribution of Romans to civilization	17	17
6	Contribution of Hebrews to civilization	17	12
CHEMISTRY			
Bell Test			
5	Two substances used in making hydrogen	16	3
10	Two gases used to bleach cloth or flowers	16	5
15	Two commercial uses of ammonia	16	0
20	Commercial uses of nitrates	16	10
25	How many grams of water are formed by the combustion of 25 grams of hydrogen in air	16	0

TABLE VII.—NUMBER OF CORRECT RESPONSES ON REPRESENTATIVE QUESTIONS IN FIVE STANDARD TESTS—(*Continued*)

TEST AND NO. OF QUESTION	STATEMENT OF THE QUESTION, LITERALLY OR SUBSTANTIALLY	TOTAL IN GROUP	TOTAL RIGHT
PHYSICS			
Starch Test			
MECHANICS			
2	Definition of erg	17	0
9	How high is a cliff if it takes a stone 5 seconds to drop to ground below?	17	1
28	Pressure of atmosphere at sea level	17	3
HEAT			
1	50 degrees Centigrade equals how many degrees Fahrenheit?	17	2
3	Definition of calorie	17	4
SOUND			
6	Velocity of sound	17	1
7	Pitch of tone depends on what?	17	2
LIGHT			
1	Meaning of complementary colors	17	3
6	Velocity of light	17	1
MAGNETISM, ELECTRICITY			
10	Definition of ampere	17	0
11	Definition of ohm	17	0
16	Definition of watt	17	0

Assuming that the standard tests used in this study are fair tests of high school subjects, it is apparent from the sample responses presented in Table VII that a great many facts learned in high school are retained after three, four, or five years by even the most intelligent of high school pupils.

If not one college senior out of a group of fifteen, sixteen, or seventeen can remember who was president during the Mexican War or can give the approximate date of the Hejira, or can define an erg, an ampere, an ohm, or a watt, and if only one, two, or three can state the historical significance of the Church Council of Nicaea or the Peloponnesian War, or can name two substances used in making hydrogen or give the velocity of light and sound, we may very well question the desirability of teaching in each subject the great mass of facts that we ordinarily teach. If we would limit our teaching in each subject to the materials possessing social worth and give each

its proper emphasis, we should undoubtedly find greater retention of subject matter after the lapse of several years. The reorganization that has been and is now going on in Mathematics, Science, and the Social Sciences promises much in the way of casting out of the courses of study the subject matter that once found a place for traditional or disciplinary reasons. Once this is done it becomes possible to take the socially valuable material remaining and, using it as a basis, construct new courses of study containing materials worth while and which we may reasonably expect pupils to remember. In time we may expect the technique of curriculum building as worked out by Professor Bobbitt and applied to Los Angeles and the principles of curriculum reconstruction as outlined by Professor Briggs to have their effect on high school curricula and courses of study.

Permanence of learning is measured not only by the amount retained after a period of time but also by the time required for relearning. It was not possible in this study to measure relearning. There is, however, a suggestion of relearning in the case of United States History. The Van Wagenen Test was used first and followed two or three weeks later by the Barr Test. The Van Wagenen tests were scored in class and considerable discussion accompanied the scoring. The Barr Test contains some elements common to the Van Wagenen. Figure 1 shows that the results on the Barr Test far exceeded the results on the Van Wagenen. It should be kept in mind, though, that the Van Wagenen Test is purely a test of information while the Barr Test is largely one involving judgment and reasoning.

It is believed by the writer that o ne of our greatest needs in Secondary Education today is a psychology of secondary school subjects. A very important part of such psychology would be concerned with the problem of retention. We need a large number of scientific studies to determine for each subject what kind of material is best retained and for the longest time, and what methods of instruction are best to insure the maximum of retention.

III. Correlations of Intelligence with Test Scores, School Marks, and Interest in School Subjects

(*A*) *Correlation of Intelligence with Test Scores.*

Subject	No. Cases	r	Table
Latin	17	.486	2
Ancient History	16	.323	11
U. S. History (Van W.)	15	.790	20
U. S. History (Barr)	15	.827	29
Geometry	18	.242	36
Physics	16	.486	45
Chemistry	16	.528	54
Average of *r*'s	..	.526	

(*B*) *Correlation of Intelligence with School Marks.*

Subject	No. Cases	r	Table
Latin	17	.000	3
Ancient History	16	−.089	12
U. S. History	15	−.054	21
Geometry	18	−.018	37
Physics	16	.000	46
Chemistry	16	−.071	55
Average of *r*'s	..	−.039	

(*C*) *Correlation of Intelligence and Interest.*

Subject	No. Cases	r	Table
Latin	17	.594	5
Ancient History	16	.000	14
U. S. History	15	.816	23
Geometry	18	.610	39
Physics	16	.662	48
Chemistry	16	.860	57
Average of *r*'s	..	.590	

It is evident that there is a substantial positive correlation between intelligence and test scores and a positive correlation between intelligence and interest in school subjects. The correlations in the latter case were computed by the Pearson cos π method. This is a very rough method as it does not take account of relative position. Furthermore, as is explained in connection with Table V it was necessary to place the *C* ratings above or below the average on a rather arbitrary basis. The correlations appearing under III *C*, then, must not be regarded as true correlations but as merely indicative of a positive or negative relationship. Table III *B* shows no correlation (−.039) between intelligence and school marks. This is just what we should expect considering the fact that the subjects of the study

come from thirty different high schools. Had they all come from the same high school we sould expect a small positive correlation. The result which we find is simply another evidence of the lack of standardization of school marks.

IV. Correlations of Test Scores with School Marks, Interest in School Subjects, Lapse of Time, and School Marks with Interest

(*A*) *Correlation of Test Scores with School Marks.*

Subject	No. Cases	*r*	Table
Latin	17	.192	4
Ancient History	16	.541	13
U. S. History (Van W.)	15	.018	22
U. S. History (Barr)	15	.192	30
Geometry	18	−.176	38
Physics	16	.275	47
Chemistry	16	.259	56
Average of *r*'s	..	.186	

(*B*) *Correlation of Test Scores with Interest.*

Subject	No. Cases	*r*	Table
Latin	17	.875	6
Ancient History	16	.409	15
U. S. History (Van W.)	15	.821	24
U. S. History (Barr)	15	.749	31
Geometry	18	.610	40
Physics	16	.860	49
Chemistry	16	.860	58
Average of *r*'s	..	.740	

(*C*) *Correlation of Test Scores with Lapse of Time Since Subjects were Studied in High Schol.*—The tabulations which deal with this matter do not lend themselves to statistical treatment for correlations. Inspection of the tables shows no relationship between test scores and time since subjects were studied. Differences of one, two, or three years do not seem to affect the score which will be made on a standard test.

Subject	No. Cases	*r*	Table
Latin	17	.875	7
Ancient History	16	.660	16
U. S. History	15	.949	25
Geometry	18	.880	41
Physics	16	.700	50
Chemistry	16	.684	59
Average of *r*'s	..	.791	

From the correlations under *IV*, *A*, *B*, *C* and *D*, it is seen that there is a rather high degree of relationship between test scores and interest, and school marks and interest but practically no correlation between test scores and school marks nor between test scores and time since studying the subjects.

V. Comparison of Sexes and Comparison of High School Subjects as to Permanence of Learning

(*A*) *Sex Differences.*—The table which follows shows sex differences in intelligence, high school marks, test scores and median interest ratings.

Subject	Average intelligence	Average marks	Average scores	Median interest	Table
Latin:					
M	166	79	58	D	1, 9
W	161	84	49	D	1, 9
Ancient History:					
M	161	83½	163	B	10, 18
W	154	89	132	C	10, 18
U. S. History:					
M	171	82	67 (Van W.) 59 (Barr)	B	19, 27
W	158	86	40½ (Van W.) 53 (Barr)	C	19, 27
Geometry:					
M	170	79	74	C	35, 41
W	160	84	44.7	C	35, 41
Physics:					
M	158	79	9.7	C	44, 52
W	164	84	13.7	C	44, 52
Chemistry:					
M	163	82	33	C	53, 61
W	157	80	20	D	53, 61

The following facts are brought out in the above table: (1) In each group the men surpass the women in intelligence as measured by our tests except the group in Physics where the women are 6 points ahead; (2) the women excel the men in school marks in all subjects except Chemistry where the men are 2 points ahead; (3) the men excel the women in all test scores except in Chemistry where the women are

4 points ahead; (4) the two sexes have the same median interest rating in Latin, Geometry and Physics while in the remaining three subjects the men are one step higher. It is interesting to note that in the subject in which the women excel the men in intelligence, namely Physics, they also excel in the test.

(*B*) *Comparison of High School Subjects as to Permanence of Learning.*

The main facts as to differences among high school subjects with respect to permanence have already been brought out in connection with Table VI and Fig. 1 on page 472. There it was seen that retention seems to be greatest in United States History, Ancient History, and Geometry and least in Chemistry, Physics, and Latin. Whether or not these six subjects would hold the same relative positions as they do now if we could make proper allowances for incidental learning is a question which cannot be answered in this paper.

VI. Summary and Conclusions

1. With college seniors, as measured by the tests in this study, retention of subject matter studied in high school but not continued later is greatest in the case of American History, second in the case of Ancient History; and so on down with Geometry, Latin, Chemistry, and Physics in the order named.

2. There is a positive correlation of about .5 between intelligence as measured by our tests and test scores.

3. There is no correlation between intelligence and school marks.

4. There is a rough correspondence between intelligence and interest in school subjects.

5. There is practically no correlation between test scores and school marks.

6. There is a positive correspondence between test scores and interest in school subjects.

7. There is no correlation between test scores and the time differences in our study.

8. There is a positive correlation between school marks and interest in school subjects.

9. According to our tests men slightly excel women in intelligence and in test scores while the women excel the men in high school marks.

A STANDARDIZATION OF CERTAIN OPPOSITES FOR CHILDREN OF GRADE SCHOOL AGE

GRACE ARTHUR

University of Minnesota

In a series of group tests[1] given in the spring of 1918 to the pupils of a grade school in St. Paul, were two lists of words used as opposites tests. Each list was composed of 20 words which I had selected from the tables of King and Gold.[2] These tables give the relative difficulty of a large number of stimulus words when used as opposites tests for educated adults. They are based upon the responses of one hundred subjects selected from the students and faculty of the University of Iowa. According to these tables, the average frequency of correct response for the first list was 99.1, and for the second list, 93.25. Hence, the first list was called the "easy opposites" test, and the second, the "hard opposites," and will be referred to as such throughout this article. The words of each list were of approximately equal difficulty, according to the tables from which they were taken, and the halves of each list balanced: that is, the first half of the list showed the same average frequency of correct response, for adults, as the second half.

When these two tests were given to grade school children and the results scored, the easy list proved to be easier as a whole than the hard list as a whole. But it also appeared that for children, certain words on the "easy" list were more difficult than some of the words on the so-called "hard" list. It seemed, too, that some words of equal difficulty for adults, as WAR, YES and ASLEEP, were of widely varying degrees of difficulty for younger subjects. Of course, it was to be expected that the difficulty of opposites as shown by frequency tables would be quite different for children from what it is for adults. Free association frequency tables compiled for children[3] differ enormously from those for adults. Moreover, those of Woodrow and Lowell show that one of the most striking differences between the free associations of the two groups is that adults have a very much stronger tendency to reply to any stimulus word with its opposite than do children.

Because of this known difference between the responses of children and those of adults in free association tests, and because of the apparent difference in quality of response in the opposites tests just described, it seemed best to make a special investigation with children in regard to opposites, before leaving the opposites tests of the group scale in

their final form. Such an investigation involves determining the difficulty of specific stimulus words for children. This may be done by presenting the stimulus words separately under uniform conditions and securing responses from a large number of subjects. The difficulty of the stimulus word is shown by the number of subjects that are unable to supply the correct opposite. The relative difficulty of the various stimulus words is determined by comparing the number of subjects failing each, the percentage of subjects failing each, or the portions of the area of the normal curve included by these percentages.

From the apparent lack of agreement between the results for adults and those for children, it was argued that a similar though less marked difference might be expected between the responses from children of the upper grades and those from the lower. To adequately measure the difficulty of a word for children, then, it must be given to children of as many degrees of ability as possible. This might be assumed to be especially true of words that were very easy or very hard. The very easy words would elicit a correct response from practically all of the upper grade children and so, as with adults, the results would be well-nigh valueless for the purpose of ranging these words in order of difficulty. Hence, to find the relative difficulty of the easier words, it would be necessary to get responses from younger subjects to whom they would offer a real test of ability. Conversely, the very hard stimulus words would be failed by all the younger pupils, and so these words would depend upon the upper age groups to determine their relative difficulty. Involved in this phase of the problem is the factor of the method of presentation and response. Means[4] used oral presentation and oral response; King and Gold made use of visual presentation and oral response; while Greene[5] preferred visual presentation and written response. For group test purposes, the last is perhaps the most satisfactory of all methods. It is evident, however, that for children especially, the difficulty of a word with visual presentation and written response may be quite different from the difficulty of the same word with oral presentation or oral response, or both. It is, in fact, possible that the difference between the frequencies of correct response for children and those for adults may be partly accounted for by reading and spelling difficulties on the part of the younger children. That it is not entirely accounted for in this way is indicated not only by the work of Woodrow and Lowell but by the difference between adult responses and those of children from the upper grades where reading and spelling difficulties have been largely overcome. As group test

conditions make visual presentation and written response the most satisfactory method for the use of opposites tests in a group scale, it is necessary to use the same method in the evaluation of stimulus words to be used later in such a scale. In this way, whatever reading and spelling factors exist may be taken into account and minimized.

It is also evident that with a list of 20 stimulus words before him, a subject may take two or three times as long in finding the exact opposite of one word as he needs for the next. To class the words as equal in difficulty is manifestly inaccurate, even though both eventually elicit a correct response. Although the time element is discussed by King and Gold, and again by Greene, it is disregarded in both investigations, in the evaluation of the stimulus words used. In both tables the difficulty of the stimulus word is expressed in terms of per cent passing, or per cent failing to find the exact opposite, without regard for the time required to do it. In the work of Means, the time element was taken into account in the giving of the stimulus word and in the calculation of a point value.

For the purpose of the present investigation it was necessary to select a number of stimulus words to be standardized; to present them to a large number of subjects of widely varying degrees of ability; to present the stimulus words visually and to require a written response; to present each stimulus word separately, and to allow a uniform amount of time for each response.

From the tables of King and Gold, I selected ninety-six words having adult frequencies ranging from 100 down to 60. In order that the actual difficulty of each word apart from all others might be determined for the children tested, each stimulus word was printed in large type on a strip of cardboard eight inches long and two inches wide. These words were arranged in four piles of 24 cards each. A sheet of ruled theme paper was given to each child. He was instructed to fold it lengthwise, and to number the lines (24) on the face uppermost. The following formula was then used: "Each of these cards has a word printed on it. When I say 'ready' you will look at the card and write as quickly as you can on your paper not this word, but a word that means JUST THE OPPOSITE. When I say 'attention' you will stop writing instantly whether you are through or not, and hold up your pencil like this. If you can't think of the opposite for any word, make a mark like this (X). If the word printed on the card were 'bad' what would you write?" After the children were shown what was meant by an opposite by means of this word, a preliminary drill was given

with the words DOWN, GO and POOR, so that every child was made to understand the mechanical part of the procedure. The first instructions were then supplemented by the following: "Now remember, you are going to write just one word after each number, and that word is to be the opposite of the word printed on the card. Ready." Each card was exposed for 10 seconds, at the end of which time "attention" was called. Thus, equal time for each word was secured. As each group of words was finished, the child turned over his paper and numbered the lines on the face then uppermost in preparation for the next group. Before the last group of words was presented the children were told that they were not expected to get them all as some of them were hard even for grown people, but that I wanted to see how many they could get.

The 96 stimulus words were presented to all the children above the first grade of a city school.[6] Thus responses were secured from 600 subjects from grades II to VIII inclusive. Twenty-five of the grade II subjects were children of more than average ability who had been in school less than a year.

The mechanical part of the work was carried out by the children with a success that was highly gratifying. Even the little grade II subjects responded promptly to the signals. The papers of four who evinced too much interest in the work of their neighbors had to be destroyed, but no other symptoms of copying were noticed either by the teachers or myself.

When it came to scoring responses, a new problem presented itself. Neither the standards of King and Gold nor those of Greene could be used without modification, owing to the much greater variety of response on the part of the children and the faulty spelling of the younger subjects in particular. The amount of latitude to be allowed in the matter of spelling had, manifestly, to be decided arbitrarily. To make the results represent something more than one person's opinion, 10 judges were secured; five psychologists and five teachers. Each judge was given a list of the stimulus words with all the responses received for each, and was asked to check each response that in his or her opinion should be given full credit, and to mark in a different way each response deserving half-credit. The judges were reminded that this was not a spelling test, but a test in the ability to summon correct opposites. They were asked, therefore, to disregard spelling as such, but to grade very strictly with regard to the exactness of the association. From the 10 scored lists thus secured, the final standards

were determined upon the following basis: in order to receive full credit, a response had to have the full credit vote of eight judges, or the equivalent (as seven full credit votes and two-half credit votes, etc.). To receive even half-credit, a response needed ten half-credit votes or their numerical equivalent, and these votes must represent at least six of the judges: *i.e.*, two whole votes and six half votes, three whole votes and four half votes, four whole and two half, five whole and one half, or six whole votes. In twenty-three cases it was found that the scoring of a response was not consistent with the rest of the scoring. This is not surprising out of a total of more than 3500 words to be scored. These inconsistencies were remedied by applying the rules which seemed to govern the rest of the scoring. Verb forms were given full credit as well as noun or adjective, if the stimulus word could be interpreted as having that form. *Vice versa*, the noun given as an opposite was as acceptable as a verb or an adjective if the stimulus word could be regarded as a noun. A change in the part of speech from that of the stimulus word in form or tense received half credit. In one case, the inconsistency lay in accepting a given spelling of a response in one case, and in rejecting the same spelling of the same word later when it was offered as the opposite of another word demanding the same response.

In Table I are given the 96 stimulus words used in the present investigation ranged in their order of difficulty for children, from the easiest to the hardest, with the percentage of frequency of correct response for each, and the point value derived from this percentage. In a parallel column are given the same words arranged in order of difficulty for adults, with their respective adult frequencies of correct response. The point values, in terms of σ, are secured by passing from the percentage of correct response to the percentage of failure for each stimulus word and reading directly from table that is based upon the area of the normal curve and assumes the base line to be broken off $\pm 3.00\ \sigma$.[7] The abilities involved in this investigation are assumed to fit this curve and the percentage of subjects who fail to give the correct opposite to a given stimulus word corresponds to the percentage of the area under the curve from the 0 point to a point on the base line, measured in units of σ. The stimulus words are numbered in order that the differences in order of difficulty of the words for the two groups can be grasped readily. YES ranks as the easiest word for both groups, and CONSERVATIVE as the hardest. The greatest amount of displacement shown by any word is 80 for BIG, which has

a rank of 8 for children and 88 for adults. This difference in rank is no doubt partly due to the methods of scoring. The word showing the next greatest difference in position for the two groups is COUNTRY with a difference in rank of 54. This difference cannot be so readily explained. Neither can the fact that IN, which ranks fourth for children, falls to the fifty-fourth place for adults. HOT, ranking seventh for children, holds for adults the fifty-fifth place. The total displacement for the 96 words is 1493, and the average displacement per word is 15.5. The correlation between the two lists, determined by the Pearson method adapted to rank differences, was found to be .74, PE = .03.

Further inspection of Table I substantiates the impression gained in scoring the responses to the lists of the original group scale. Of the words of the original "hard" list, half are easier than WAR on the "easy" list. Of the "easy" list, only two words have as high a frequency of correct response for children as OUTSIDE on the "hard" list. For adults, QUICK, SOMETHING and LOVE are the three easiest words on the "hard" list, with a frequency of 95 per cent. For children, QUICK ranks fourth in difficulty of the words of this list, SOMETHING ranks eighth, and LOVE ranks eleventh.

Greene's list was selected from the Hard Stimulus Words of King and Gold. Thirty-two words of his list appear also in the present investigation. A comparison of the ranking of these words according to his results with their ranking for children yields a total displacement of 173, an average displacement of 5.4 per word, and a correlation of .75, PE = .05.

Arguing from these comparisons of the responses of the 100 educated adults of King and Gold and those of Greene's larger group of college freshmen with the responses of 600 children of a given grade school, it would seem that the ability to give opposites differs in children from that of adults not only quantitatively but qualitatively; that the children are no more "adults in miniature" in regard to this ability than they are in other respects; and that the ease or difficulty of a test of this ability for adults gives only a very rough index of its ease or difficulty for children. What factors constitute this ease or difficulty, it is not the purpose of this investigation to determine. It would be quite possible to eliminate the factor of group reaction by presenting the same stimulus words to the same subjects in the same manner, individually, and to eliminate reading and spelling by using a different method of presentation and response. It might be worth

while, also, to construct a table showing the school grade in which the various stimulus words reach adult frequency. For instance, SUMMER, with an adult frequency of 100 per cent reaches the same frequency of correct response among grade V subjects, while CONSERVATIVE, with an adult frequency of 60, has for even grade VIII subjects a frequency of only 1.04 per cent.

As stated at the outset, the reason for this investigation was the need for definite information as to the relative difficulty of certain stimulus words when used as opposites tests for children, in order that lists of words of known difficulty might be constructed for use in a group scale. It is evident that Table I furnishes the desired data, and that from it such lists can be prepared of hard or easy stimulus words, and of words of equal or graded difficulty, according to the purpose of the test.

It is interesting to compare the lists of opposites used by other investigators with the results shown by this table. Whipple[8] publishes three lists of "easy opposites" prepared by Woodworth and Wells. "Lists I and II . . . are presumed to be of equal difficulty and to be so arranged that the last half is just as difficult as the first half." List III, which is the list of easy opposites recommended by Whipple, "is a selection of the 20 easiest opposites in Lists I and II." Fourteen words of List I, thirteen of List II and fourteen of List III were used as stimulus words in the present investigation. Those from List I with their respective frequencies for children are: soft 85.08, white 80.58, up 88.16, dead 76.66, hot 87.75, asleep 80.33, wet 86.50, high 86.66, dirty 83.66, east 85.50, day 82.83, yes 92.08, wrong 82.82, empty 81.00.

The thirteen from List II are: north 81.83, sour 71.08, weak 66.00, good 84.00, after 63.91, above 76.25, sick 74.25, rich 91.76, love 61.50, tall 52.75, open 74.58, new 85.16, come 86.58.

Those from List III are: high 86.66, white 80.58, yes 92.08, above 76.25, north 81.83, wet 86.50, good 84.00, rich 91.76, up 88.16, hot 87.75, east 85.50 day 82.83, big 87.76, love 61.50.

As the lists are not complete, and as there is no way of making allowance for possible differences in methods of procedure and scoring between this and other investigations, comparisons are of necessity only suggestive and cannot be made the basis of any definite conclusions. But it would seem that when used as group tests for children, Lists I and II are far from being of equal difficulty, and List III includes some fairly difficult opposites and omits many of the easiest.

For the words here given, List I shows a range in frequency of correct response for children of 15.42 per cent, from 92.08 to 76.66; List II, according to Whipple, "of presumably equal difficulty," ranges from 91.76 to 52.75, which is a range of 39.01 per cent. List III, made up of the "easiest opposites from Lists I and II" has a range of 30.58, or from 92.08 to 61.50 per cent.

As a substitute for List III, I would submit the following list of 20 stimulus words that according to Table I are among the easiest for children: rich 91.76, outside 90.83, in 89.16, young 88.75, up 88.16, hot 87.75, big 87.66, high 86.66, come 86.58, wet 86.50, winter 85.75, east 85.50, new 85.16, soft 85.08, brother 85.00, good 84.00, dirty 83.66, easy 83.16, day 82.83, wrong 82.83.

This list has the advantage of being made up of words that vary but slightly from each other in difficulty. The range in frequency of correct response is only 8.93 per cent as compared with the range of at least 30.58 of "List III." The words are presented in their order of difficulty, but can easily be arranged to give a list in which the halves and quarters are balanced.

If, instead of the above list of easy words of equal difficulty, parallel lists of easy words of gradually increasing difficulty are needed, the following can be used:

young	88.75	in	89.16
hot	87.75	up	88.16
high	86.66	big	87.66
come	86.58	wet	86.50
east	85.50	winter	85.75
new	85.16	soft	85.08
good	84.00	brother	85.00
easy	83.16	dirty	83.66
wrong	82.83	day	82.83
empty	81.00	north	81.83
quick	80.58	white	80.58
near	79.58	asleep	80.33
dead	76.66	heavy	77.16
open	74.58	above	76.25
late	73.75	sick	74.25
something	72.83	first	73.25
sour	71.08	sharp	72.00
country	68.33	stay	69.50
here	68.08	land	68.08
weak	66.00	sell	67.41

The range for these 40 stimulus words is only 23.16 per cent and this again is less than that of Whipple's easiest 20. Balanced lists,

as contrasted with graded lists, made up from them, would probably be of more nearly equal difficulty than those in use for children up to the present time.

For use in the group scale, the lists of Table II were arranged. They begin with the very easiest words of Table I and grade upward by approximately equal and parallel steps to the hardest. This arrangement tends to eliminate the erratic results that sometimes follow when children are discouraged by the difficulty of hard stimulus words that appear near the top of a list, and so never reach the easier words farther down with which they could cope successfully. By the use of these lists, equal time can be allowed for all grades from the second to the twelfth, inclusive. The score can be determined by simply counting the number of correct responses, or by adding their point values. The use of parallel lists tends to increase the reliability of the final score. These lists should afford an adequate measure of the ability of a child to give opposites, as they are based upon the performance of other children.

That these words had been adequately standardized for children was demonstrated two years later, when, in the spring of 1921, the lists of Table II were incorporated in the group intelligence scale above referred to and given to all the pupils of Chisholm above grade I. An interval of one month was allowed to elapse between the giving of the two series. The following formula was used: "This paper has some words printed on it. I want to see how quickly you can write a word after each of these words that will mean just the opposite. If the word on the paper were 'cold' you would write . . .? If the word on the paper were 'little' you would write . . .? and so on. (BAD and WHITE were also used as additional illustrations where needed). When I say 'ready,' take your pencil in the hand you write with, and take hold of the corner of the paper with your other hand. When I say 'turn,' turn your paper over and begin to work as fast as you can. Keep on working until I say 'stop.' But the instant I say 'stop,' put your pencil down. Ready. Turn."

Equal time, 90 seconds, was allowed all groups. To insure uniformity of conditions, I gave all tests in person. Norms were determined for age groups 6 to 18, inclusive, for both lists. The 17 and 18-year age norms might be modified by the results from additional cases, as the latter especially appears to be a somewhat selected group. Table III presents these norms together with the number of cases for each age group and the standard deviation. In scoring, the

TABLE I[9].—STIMULUS WORDS IN ORDER OF DIFFICULTY FOR CHILDREN, WITH PERCENTAGES OF CORRECT RESPONSE FOR EACH, AND POINT VALUE IN TERMS OF σ; AND THE SAME WORDS ARRANGED IN ORDER OF DIFFICULTY FOR ADULTS, ACCORDING TO THE TABLES OF KING AND GOLD

Stimulus	Children's Fahrenheit	Point value	Stimulus	Adult's Fahrenheit
1 yes	92.08	1.59	yes	100.0
2 rich	91.75	1.62	young	100.0
3 outside	90.83	1.67	wet	100.0
4 in	89.16	1.77	winter	100.0
5 young	88.75	1.79	east	100.0
6 up	88.16	1.82	dirty	100.0
7 hot	87.75	1.84	asleep	100.0
8 big	87.66	1.84	war	100.0
9 high	86.66	1.89	remember	100.0
10 come	86.58	1.89	new	99.0
11 wet	86.50	1.90	brother	99.0
12 winter	85.75	1.93	day	99.0
13 east	85.50	1.94	north	99.0
14 new	85.16	1.96	dead	99.0
15 soft	85.08	1.96	open	99.0
16 brother	85.00	1.96	wild	99.0
17 good	84.00	2.01	friend	99.0
18 dirty	83.66	2.02	after	99.0
19 easy	83.16	2.04	sell	99.0
20 day	82.83	2.05	rich	98.0
21 wrong	82.83	2.05	high	98.0
22 north	81.83	2.09	come	98.0
23 empty	81.00	2.12	good	98.0
24 white	80.58	2.14	empty	98.0
25 quick	80.58	2.14	white	98.0
26 asleep	80.33	2.15	heavy	98.0
27 near	79.58	2.17	sharp	98.0
28 heavy	77.16	2.26	sour	98.0
29 dead	76.66	2.27	straight	98.0
30 above	76.25	2.29	soft	97.0
31 open	74.58	2.34	easy	97.0
32 sick	74.25	2.35	to win	97.0
33 late	73.75	2.36	wrong	96.0
34 first	73.25	2.38	weak	96.0
35 something	72.83	2.39	up	95.0
36 sharp	72.00	2.42	quick	95.0
37 sour	71.08	2.44	something	95.0
38 stay	69.50	2.49	love	95.0
39 country	68.33	2.52	beautiful	95.0
40 land	68.08	2.53	outside	94.5
41 here	68.08	2.53	here	94.5
42 sell	67.41	2.55	above	94.0
43 weak	66.00	2.59	buy	94.0
44 wild	65.91	2.59	backwards	94.0
45 remember	65.00	2.61	forget	94.0
46 after	63.91	2.64	increase	94.0

TABLE I.—(*Continued*)

Stimulus	Children's Fahrenheit	Point Value	Stimulus	Adult's Fahrenheit
47 straight	62.83	2.67	rough	94.0
48 war	62.50	2.68	cheap	93.5
49 love	61.50	2.71	near	93.0
50 less	61.16	2.72	late	93.0
51 nowhere	58.83	2.78	broad	93.0
52 absent	58.33	2.79	many	93.0
53 win	57.58	2.81	succeed	93.0
54 beautiful	55.41	2.86	in	92.5
55 best	53.75	2.90	hot	92.0
56 buy	53.00	2.92	strength	92.0
57 tall	52.75	2.93	left	91.5
58 backwards	52.33	2.94	tall	90.5
59 left	49.33	3.01	beginning	90.5
60 over	48.16	3.04	land	90.0
61 friend	45.75	3.10	stay	89.0
62 beginning	44.08	3.14	absent	89.0
63 forget	43.50	3.16	less	88.0
64 stale	42.66	3.18	nowhere	87.0
65 rough	41.75	3.20	over	87.0
66 push	41.66	3.20	generous	87.0
67 broad	41.41	3.21	bless	87.0
68 raise	40.16	3.24	like	86.0
69 lack	39.16	3.27	part	86.0
70 cheap	37.00	3.32	simple	86.0
71 many	36.50	3.34	sick	85.0
72 much	35.00	3.38	lack	85.0
73 increase	34.75	3.38	to take	85.0
74 float	34.25	3.40	refined	85.0
75 take	34.08	3.40	false	84.0
76 innocent	32.00	3.46	stale	83.0
77 tender	29.33	3.54	raise	83.0
78 strength	27.25	3.60	to float	82.0
79 false	23.33	3.72	vertical	81.0
80 generous	22.75	3.74	silly	81.0
81 rude	19.25	3.86	first	79.0
82 like	17.66	3.92	best	78.0
83 succeed	15.50	4.00	push	77.5
84 vertical	15.25	4.02	tender	76.0
85 stormy	14.08	4.07	past	76.0
86 part	13.00	4.12	injurious	76.0
87 past	9.83	4.28	innocent	75.0
88 bless	7.91	4.40	big	72.0
89 silly	7.50	4.43	permit	69.0
90 frequently	4.58	4.67	rude	67.0
91 refined	4.16	4.71	much	66.0
92 injurious	3.16	4.83	frequently	66.0
93 permit	2.16	4.99	country	65.0
94 strict	.83	5.33	stormy	64.0
95 simple	.83	5.33	strict	62.0
96 conservative	.16	5.75	conservative	60.0

point value of a word was disregarded and the score was simply the sum of the full credit and half-credit responses. From these results it would appear that the lists were, for the children of Chisholm, of practically equal difficulty. Below the 17-year age group, the largest difference between the norms for the two lists at any age level is .62. At the 17-year age level, the difference is 1.11, but this may be due to the limited number of cases at this point, as the difference between the norms for the two lists at the 18-year level again drops to .50.

The standardization of a large number of additional stimulus words for children by the same method as that described above, would make possible the arrangement of new tests as alternatives to those here presented, and of others of various degrees of difficulty as the need for them in the practical use of opposites tests becomes evident.

TABLE II.—PARALLEL LISTS OF GRADED DIFFICULTY, BASED UPON CHILDREN'S FREQUENCY TABLES

Stimulus	Per cent passing	Point value	Stimulus	Per cent passing	Point value
yes	92.08	1.59	rich	91.76	1.62
in	89.16	1.77	young	88.75	1.79
new	85.16	1.96	east	85.50	1.94
north	81.83	2.09	empty	81.00	2.12
heavy	77.16	2.26	dead	76.66	2.27
sharp	72.00	2.42	sour	71.08	2.44
weak	66.00	2.59	wild	65.91	2.59
love	61.50	2.71	less	61.16	2.72
best	53.75	2.90	buy	53.00	2.92
over	48.16	3.04	friend	45.75	3.10
push	41.66	3.20	broad	41.41	3.21
much	35.00	3.38	increase	34.75	3.38
tender	29.33	3.54	strength	27.25	3.60
false	23.33	3.72	generous	22.75	3.74
rude	19.25	3.86	like	17.66	3.92
stormy	14.08	4.07	part	13.00	4.12
bless	7.91	4.40	silly	7.50	4.43
frequently	4.58	4.67	refined	4.16	4.71
injurious	3.16	4.83	permit	2.16	4.99
strict	.83	5.33	simple	.83	5.33

TABLE III.—NORMS FOR PARALLEL LISTS OF GRADED DIFFICULTY

Chronological age	Cases	Average	σ	Cases	Average	σ
6	232	.04	.31	222	.05	.37
7	294	1.11	1.74	295	1.29	2.00
8	259	2.83	2.45	256	3.45	2.74
9	272	4.27	2.97	278	4.79	2.71
10	223	6.11	3.09	217	6.47	2.49
11	239	7.25	3.35	240	7.54	2.41
12	212	8.51	3.32	218	8.58	2.74
13	187	10.62	3.20	191	10.23	2.76
14	163	11.07	3.28	166	10.93	3.21
15	119	11.87	3.48	122	12.27	3.26
16	70	13.13	3.02	66	13.72	2.26
17	55	13.81	2.84	57	14.92	2.61
18	24	15.87	1.92	27	15.37	1.63

[1] Arthur and Woodrow: An Absolute Intelligence Scale. *Journal of Applied Psychology*, Vol. III, 1919, pp. 118–137.

[2] King and Gold: A Tentative Standardization of Certain Opposites Tests. *Journal of Educational Psychology*, Vol. VII, 1916, pp. 459–482.

[3] Woodrow and Lowell: "Children's Association Frequency Tables," Psychological Monographs No. 97, 1916, pp. 81–82.

[4] Means, Marie Hackl: "A Tentative Standardization of A Hard Opposites Test." Psychological Monographs No. 134, pp. 22.

[5] Greene, Harry A.: A Standardization of Certain Opposites Tests. *Journal of Educational Psychology*, Vol. IX, 1918, pp. 559–566.

[6] Grateful acknowledgment is due Miss Edith Taylor, principal of the Hancock School, St. Paul, Minnesota, where this part of the work was done, and to the teachers for their hearty cooperation.

[7] Rugg, H. O.: "Statistical Methods Applied to Education." Houghton Mifflin Company, 1917, Table VI, pp. 396–400.

[8] Whipple, Guy M.: "Manual of Mental and Physical Tests." Warwick and York, 1915, Vol. II, pp. 79.

[9] Tabulation on which this material is based, with responses in full for the 600 cases, is on file in the library of the University of Minnesota.

FURTHER NOTE UPON THE PROBABLE ERROR OF THE MEAN

TRUMAN L. KELLEY

Stanford University

In Dr. Holzinger's reply[1] to my note[2] upon his formula (6)[3] for the probable error of the mean, occur the following statements:

(*a*) "I therefore suggest that formula (1)[4] be held in abeyance until further evidence is at hand."

(*b*) "I do not recommend the use of formula (2)[5] as a substitute because, even though more reasonable than (1) assumptions as to the correlation of errors are still not clearly justified."

(*c*) (After a brief development)" . . formula (6) follows at once, since ${}_1\sigma_x = \sigma_x$. Thus Professor Kelley's reasoning may lead him to the formula to which he is objecting."

As I cannot at all agree to the equation

$$\sigma_x{}^{2'} = \sigma_x{}^2{}_1 + \sigma_2$$

and further cannot follow Dr. Holzinger's reasoning where he states "formula (6) follows at once," I see no chance that my reasoning would lead me to formula (6). However, since Dr. Holzinger has withdrawn this formula there is no point in further defending my objection to it.

Dr. Holzinger introduces in his reply to my note the following formula (2):

$$\sigma_x{}^{2'} = {}_1\sigma^2 r_{12}$$

and raises a question as to the authorship of it which it is incumbent upon me to note. When I first derived formula (2)[6] I was unaware that Spearman had reached the same result. Spearman's formula for the standard deviations of sums and the formula for the correlation between sums and differences[7] have many very important applications

---

[1] *Journal Educational Psychology*, September, 1923.

[2] *Journal Educational Psychology*, September, 1923.

[3] *Journal Educational Psychology*, May, 1923.

[4] $\sigma_x{}^{2'} = \sigma_x{}^2{}_1(2 - r_{12})$.

[5] $\sigma_x{}^2 = \sigma_1 r_{12}$.

[6] *Journal of Educational Psychology*, November, 1919 (not *Journal Educational Research*, May, 1921).

[7] *British Journal of Psychology*, vol. 5, 1913.

and I am glad to have my attention called to the fact that the important formula (2) is one of the applications which Spearman noted. Let me at this time further express my regret that when I derived formulas for the standard deviations of sums and for the correlations between sums[1] I was unfamiliar with Spearman's work[2] and therefore did not credit him with the formulas which I gave.

Dr. Holzinger's statement (*b*) quoted above has no bearing which I can ascertain upon the question of the probable error of the mean, but I am glad formula (2) was mentioned as it gives me the opportunity for expressing my high regard for Spearman's contributions and my regret that upon at least two occasions I have not duly credited him.

---

[1] *British Journal of Psychology*, Vol. 5, 1923.

[2] Tables, etc. *Bulletin* of the University of Texas, No. 27, May, 1916.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY CECILE COLLOTON

Department of Educational Psychology, The Lincoln School of Teachers College

Intelligence Tests

The Use of Intelligence Tests in Educational Administration in the Providence Schools. Richard D. Allen. School and Society, 1923, Sept. 22, 335–339. Describes the changes in classification and promotion in a school system where the intelligence test is wisely used.

Tests of Candidates for the Rhodes Scholarship. J. B. Miner. School and Society, 1923, Sept. 8, 297–300. Reports a highly satisfactory use of intelligence tests as supplementary evidence to personal histories and interviews in choosing a Rhodes Scholar.

A Study of the Data on the Results Gathered from Repeated Mental Examinations of 200 Defective Children Attending Special Schools over a Period of Eight Years. Meta L. Anderson. Journal of Applied Psychology, 1923, March, 54–64. Discusses six types of mental growth curves of feeble minded children and summarizes variations in IQ.

The Freshman: Thorndike College Entrance Tests, First Semester Grades, Binet Tests. W. I. Root. Journal of Applied Psychology, 1923, March, 77–92. Reports a study of 569 Freshmen of the College of the University of Pittsburgh with special emphasis on the prognostic value of the Thorndike test.

A Mental Survey of 59 Normal School Students: Some Correlations and Criticisms. Lawrence A. Averill. Journal of Educational Research, 1923, April, 331–337. Otis Group Examination Scale, Thurstone Psychological Examination, analysis of scholastic grades, and instructors' estimates of native ability supply the data reported in this study.

A Study of the Binet and Terman Intelligence Tests with Eleven-year-old Children. George T. Avery. Journal of Educational Research, 1923, May, 429–433. Reports the Terman Group of mental ability as a reliable instrument for classification. Shows correlation of .80 with the Stanford-Binet.

Measurement of the Effectiveness of Differentiation of High School Pupils on the Basis of the Army Intelligence Tests. Clifford Woody. Journal of Educational Research, 1923, May, 397–409. From the results of a study of the scores of 83 high school freshmen on the army alpha Intelligence Test and four educational tests the author states that any good intelligence test can be used satisfactorily for grouping students of homogeneous abilities.

Some Uses of the Freshman Test in the Smaller College. School and Society, April 14, 416–417. The Thurstone Intelligence Test used with 450 subjects at Sweet Briar College over a period of three college grades does have predictive value and gives information of value to the administrator and the teacher.

Grading and Promotion. Joseph S. Taylor. School and Society, 1923, April 14, 405–409. The limitations of intelligence tests and their value in the study of individual capacities.

The Stanford-Binet Tests in Some English Schools. P. L. Gray and R. E. Marsden. Journal of Educational Research, 1923, September, 150–155. Reports the satisfactory use of the Stanford-Binet in English schools. Suggests a few changes in wording and in location of separate tests.

Relation of General Intelligence to the Persistence of Educational and Vocational Plans of High School Pupils. William Martin Proctor and Helen Ward. 1923, April, 277–288. Reports the second check-up on 771 high school pupils whose vocational ambitions, educational plans, school success, and general intelligence scores had been secured in 1917–1918.

The Reliability of MA and *IQ Based on Group Tests of General Mental Ability.* Arthur I. Gates. Journal of Applied Psychology, 1923, March, 93–100. Shows the great variation in MA and IQ computed from different group tests and warns against such a use of group test scores.

Democracy, Determinism and the IQ. John C. Almack, James F. Bursch, and James C. DeVoss. School and Society, 1923, Sept. 8, 292–295. Answers some of the arguments of the opponents of intelligence testing.

Again: Educational Determinism. Truman L. Kelley. Journal of Educational Research, 1923, June, 10–19. An analytical discussion of the main points of disagreement between Bagley and Terman.

Transmutation of Scores between Binet Tests and Group Tests. W. I. Root. Journal of Educational Research, 1923, April, 338–341. Five tables of transmuted scores of value to the administrator for rough comparison.

Educational Tests

The Briggs Form Test in Use. C. C. Certain. The English Journal, 1923, April, 244–257. Shows how publicity given to group and individual test scores aids in definite improvement. Explains procedure and forms required for the test.

Are Your Pupils up to Standard in Composition? C. C. Certain. The English Journal, 1923, June. Presents typical procedures for the use of standard tests in composition in the high school and explains how to interpret results.

Why Not Include Standard Tests in Your Teaching Program This Year? C. C. Certain. The English Journal, 1923, September, 463–480. Outlines a definite schedule of standard tests for use in English classes and gives detailed directions for scoring papers, and tabulating and interpreting results.

A Survey of the Test Movement in History. Paul Tyler Kepner. Journal of Educational Research, 1923, April, 309–325. Presents "(1) the methodology of constructing tests; (2) the types of tests in the field and their nature; (3) the criticism, destructive and constructive, passed upon the various tests."

A Study of the Validity of the Courtis and Studebaker Practice Tests in the Fundamentals of Arithmetic. W. J. Osburn. Journal of Educational Research, 1923, September, 93–105. Shows that each test covers less than two-thirds of the total field and affords too little practice on particular difficulties such as zero combinations, division with remainders, etc.

The Use of American Tests to Measure English Teaching in China. Sterling G. Brinkley. Journal of Educational Research, 1923, September, 136–144. Sum-

marizes the results of English, reading, language and spelling scales used in a typical college preparatory middle school of East China. Raises questions for the modern language teacher.

Years in School and Achievements in Reading and Arithmetic. Chas. L. Harlan. Journal of Educational Research, 1923, September, 145–149. Results from the Monroe Silent Reading Tests and the Monroe Reasoning Tests in Arithmetic show that it is not the *amount* of time spent, but *how* and *by whom* the time is spent that counts for achievement.

Science of Education. William A. McCall. Journal of Educational Research, 1923, May, 384–396. Describes the construction of an arithmetic test for use in China.

Miscellaneous

Training for Leadership. Garry C. Meyers. School and Society, 1923, April 21, 437–439. Argues for the development of sense of social obligation and traits of sympathy and likeableness in gifted children.

The Man of Genius. S. C. Kohs. School and Society, 1923, April 21, 429–431. Pleads for the moral education of the child or superior intelligence.

From Him That Hath Brains. Louis A. Pechstein. Educational Review. 1923, June, 4–6. Uses data from both grade children and college students to show how the IQ, EQ and AQ operate in measuring the educational efficiency of the *individual.*

A Study of Ten Gifted Children Whose School Progress Was Unsatisfactory. Dorothy Van Alstyne. Journal of Educational Research, 1923, September, 1923, 122–135. Reports the various means used to secure individual analysis, and the remedial measures used to bring each child up to the level of his ability. Two typical cases are given in detail.

Some Characteristics of Leadership. L. Ruth Nutting. School and Society, 1923, Sept. 29, 387–390. A study of the reasons given for choice of captains in seventh and eighth grade girls, gymnasium classes shows good judgment as to the qualities of importance to a leader. Qualities are listed. Comparisons with intelligence and popularity ratings are made.

Mental Attitude of Children towards School Work. Arthur W. Kallom. School and Society, 1923, Sept. 22, 339–343. Emphasizes the necessity for following up test results by intensive study of individuals who are exceptional in any way.

The Study of Individual Abilities—The Normal Child. Walter F. Dearborn. School and Society, 1923, Sept. 22, 331–335. Repeated tests—both mental and physical—of the same children over a period of 10 to 12 years will answer many questions about individual development. Such an experiment has been begun at Harvard.

Measuring Interest Objectively. Harold E. Burtt. School and Society. 1923, April 21, 444–448. Describes four tests designed to measure interest and industry used with 43 students specializing in agricultural engineering at Ohio State University. Results are interesting though not conclusive.

The Downey Will Temperament Group Test: Analysis of Its Reliability and Validity. G. M. Ruch and M. C. Del Manzo. Journal of Applied Psychology, 1923, 65–78. Seven tables report the statistical analysis of the scores of 146 high school students tested with the Downey Group Will-Temperament Test.

Six detailed conclusions summarize the most important findings and emphasize the need for a more valid instrument for measuring volitional traits.

Racial Differences as Measured by the Downey Will-Temperament Test. John F. McFadden and J. F. Dashiell. Journal of Applied Psychology, 1923, March, 30–53. Six tables, six figures, and a summary give the results and conclusions of a study of the comparison of temperament in individuals of the white and negro races.

An Experiment in Testing Engine Lathe Aptitude. Everett F. Patten. Journal of Applied Psychology, 1923, March, 16–29. Describes the analysis of engine lathe ability by observation of men at work and the selection of tests to measure this ability.

Herbert: A Study of Difficulty in Spelling and Reading. Bernice Leland. Journal of Educational Research, 1923, June, 49–58. A case study of special disabilities and report of remedial measures employed.

New Thoughts about the Feeble-Minded. Edgar A. Doll. Journal of Educational Research, 1923, June, 31–48. Emphasizes the importance of the special class in the social and industrial training of the feeble minded.

Another Criticism of Tests Requiring Alternative Responses. C. W. Odell. Journal of Educational Research, 1923, April, 326–330. Advantages of True-false Tests emphasized.

A Critical Study of the Right Minus Wrong Method. Paul V. West. Journal of Educational Research. 1923, June, 1–9. Shows data to prove the unreliability of the right minus wrong method of scoring.

The Effect of First-year Latin upon Kmowledge of English Words of Latin Derivation. E. L. Thorndike and G. J. Ruger. School and Society, 1923, Sept. 1, 260–270. Results from the Carr English Vocabulary Test used in ninth grade classes in schools representing 19 states show that Latin students made much greater than non-Latin students in Latin derived words. Twelve tables give statistical data.

An Individual Record Card for Preserving Test Data. Grace Arthur. School and Society, 1923, Sept. 22, 356–357. Describes a single card used in Chisholm, Minnesota, for recording standard tests results.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY LAURA ZIRBES[1]

1. *A Book on Special Abilities and Disabilities.*—For a generation data have been accumulating concerning special abilities and defects and their relation to general mental ability. Dr. Leta S. Hollingworth has done a helpful bit of work in collecting and interpreting the evidence in her recent book.[2]

It is not a detailed case-book like Dr. Bronner's "Special Ability and Disability," or a report of new investigations like Dr. Hollingworth's earlier monograph, "The Psychology of Special Disability in Spelling," or of Dr. Gates's "Psychology of Reading and Spelling." It is an *épitome* and evaluation of both theories and the data of investigations.

For example, the author sets forth certain current positions on the relationship between "capacities" (throughout the book the term capacities is used to mean *present abilities*, not merely *native capacities*) by contrasting the Spearman and the Thorndike points of view. By a review of the data of correlation she discusses the issue of *general* intelligence vs. specific abilities. The helpfulness of her book to the student of education and of educational psychology is well illustrated by the way she brings the literature of such problems up to date. The book is an inductive exhibit of the contemporary dynamic psychology. It uses throughout the data of mental and educational measurement and emphasizes the specificness of mental activity.

Special talents and defects mean to Dr. Hollingworth, those abilities demanded of an individual for success in particular school subjects—reading, spelling, arithmetic, music, and drawing. Abilities involved in other school subjects such as language (composition and grammar) and history and related social subjects are not discussed. The author's treatment of the neural basis of mental life may be summarized fairly by this sentence: "Experimental neurology has

[1] All unsigned reviews were prepared by Laura Zirbes.

[2] "Special Talents and Defects." New York, Macmillan, 1923, pp. xix–216.

nothing secure to offer by way of establishing the neural basis of the special talents and defects."

The chapter on the psychology of reading shows the high positive correlation between reading ability and the IQ, gives a three-page *resumé* of the evidence on the mechanics of reading, devotes 19 pages to reporting a new four-year study of a non-reader, and presents an interpretation of studies of both non-readers and of cases of special ability in reading. At this point in the book it seems to me there is far too little attention given to the recent studies from the Judd-Buswell laboratory.

An excellent chapter on the psychology of spelling consists of an analysis of the process of learning to spell and an *épitome* of the results of psychological examination of poor spellers. The author takes up such questions as "Can special defect in spelling be overcome?" and "Does reading teach spelling?"

A very brief chapter on arithmetic summarizes the data on such problems as the relation between IV and ability in arithmetic, mental functions employed in arithmetical calculation, the organization of arithmetical abilities, psychological studies of deficiency in arithmetic, and arithmetical prodigies. (The last named is given more space than is devoted to all the preceding matters.) Much of this chapter consists of abstracts and quotations from Thorndike and those associated with him in the investigation of arithmetic.

The author treats music and drawing in much the same way and completes her survey by a brief treatment of individuality and education. Certain miscellaneous special talents are briefly discussed, *e.g.* mechanical ability as measured by tests and important social qualities. The inadequacy of data in these two fields, especially in the latter, is revealed by the discussion of ability to lead and handle people. The data of two studies which I have made lead to the tentative conclusion that there is little correlation between such abilities and *verbal* intelligence. This is quite at variance with Dr. Hollingworth's point of view.

Her book will be very helpful to teachers who wish a brief *resumé* and interpretation of the psychology of the school subjects.

HAROLD RUGG.

---

2. *A Strategic Position in Education.*—"As is the principal so is the school" is the fundamental theme of a recent textbook[1] on the organiza-

[1] Cubberley, E. P.: "The Principal and His School." New York, Houghton Mifflin Co., 1923, pp. 570.

tion, administration and supervision of instruction in an elementary school. For two reasons the appearance of this book is particularly opportune. The first is because of the substantiating evidence which it adds to the rapidly growing conviction that the hitherto neglected and submerged elementary principal is in reality a very important person in determining the educational tone and general effectiveness of our schools. The second is the wealth of professional training material which this book makes available in a field where such material has been noticeably meagre.

The author has so filled his 570 pages with sound educational theory; pertinent principles of school administration; interesting, concrete and helpful illustrations; and suggestive and stimulating problems that the book will commend itself to a wider variety of readers than the title alone would indicate. No principal, experienced or inexperienced can read it without professional improvement resulting from every chapter. No teacher, regardless of how long he has taught, can read it, particularly the chapters on supervision, without receiving innumerable helpful suggestions and an increased professional zeal which is certain to result in more intelligent and sympathetic cooperation with the principal of his building. No superintendent or supervisor can read the clear-cut analysis of the relationship of the principal to these officers without a better understanding of this troublesome administrative problem. Chapter XVIII alone, on "Knowing the School" would justify the addition of this volume to the professional library of any principal or teacher.

The technique of scientific experimentation in education is yearly becoming more generally understood and appreciated. The number of varying theories of education and proposed methods of teaching is yearly becoming more numerous and more confusing. The conviction is growing that many schools will have to apply the test of scientific experimentation to these conflicting proposals in order to evaluate them. If this is to be done, the building principal is in the most strategic position to become a director and guide for such experiments as integral parts in his larger program of supervision and teacher growth. This phase of the principal's work seems to be the single omission of any importance in the comprehensive treatment presented in this very readable textbook.

E. S. Evenden.

Teachers College,
Columbia University.

3. *A Significant Committee Report. The Reorganization of Mathematics in Secondary Education.*—This is the most valuable report[1] that has ever been made in the field of secondary mathematics in this country. It should prove to be epoch-making in its effect, not only because of its intrinsic worth, but also because it should stimulate many individual teachers of mathematics and organizations of such teachers to carry on and extend the investigations so well begun.

The members of this committee were highly trained specialists in their various fields, either in secondary or higher mathematics. They were so chosen as to give all sections of the country fair representation in the report.

The report carries weight especially because, contrary to the usual custom, the work of this committee was subsidized. Thus, it was possible for all members to attend its sessions, and to have a hand in making the final report.

Moreover, the findings of the committee have been influenced by suggestions and help from organizations of mathematics teachers from all parts of the country. The report therefore represents in a broad sense the best thinking on the reorganization of mathematics in the United States.

The report is divided into two parts as follows: Part I, General Principles and Recommendations. Part II, Investigations conducted for the Committee.

Part I consists of eight chapters dealing with a brief outline of the report, the aims of mathematical instruction—general principles, mathematics for years seven, eight, and nine, mathematics for years ten, eleven, and twelve, college entrance requirements, a list of propositions in plane and solid geometry, the function concept in secondary school mathematics, and terms and symbols in elementary mathematics. It is this part of the report that is likely to prove most helpful to the classroom teacher of mathematics.

Part II also consists of eight chapters dealing with the present status of disciplinary values in education, the theory of correlation applied to school grades, mathematical curricula in foreign countries, experimental courses in mathematics, standardized tests in mathematics for secondary schools, the training of teachers of mathematics,

---

[1] A Report by the National Committee on Mathematical Requirements under the Auspices of the Mathematical Association of America, Inc., pp. x + 652. (Obtainable from J. W. Young, Hanover, N. H.)

certain questionnaire investigations, and a bibliography on the teaching of mathematics.

While Part II of the report may not be so generally helpful to the classroom teacher, certain parts of it are sure to affect his teaching practice and his thinking.

It may be regretted by some that it was not convenient to carry the investigations to the stage of "scientific accuracy." To answer some of the questions raised in the report will take years of careful study and further investigation. It is doubtful whether it would have been possible to hold such a group together long enough to complete a work of this type. It would have involved too great personal sacrifice.

The committee has indicated what the child should be taught, with some limits as to organization and methods. It still remains for some one to show what the child is really able to learn with profit in a given grade. Such a plan is far reaching and must necessarily take into account the entire question of method as well as content.

This report has made a noteworthy contribution to the field of secondary mathematics. The individual members of the committee deserve the thanks and congratulations of all teachers of mathematics. It is to be hoped that the material of this report may be made available to all teachers of mathematics in the secondary schools.

W. D. REEVE.

---

4. *A Textbook for College Courses.*—This is a brief, concise treatment of the intelligence testing movement. The book[1] is divided into three parts: Part I, historical and theoretical; Part II, the methods; and Part III, the results.

The book acquaints the reader with the historical background of the intelligence test and the theoretical assumptions underlying the work. It makes him familiar with typical tests—both individual and group—their general make-up, the uses for which they are best adapted, and the sources from which they can be obtained. In the final section are brought together the most important investigations that have been made in various fields. Separate chapters are devoted to each type of individual studied, such as "The Blind, The Feebleminded, The Superior, etc."

[1] Pintner, R.: "Intelligence Testing." New York, Henry Holt and Company, 1923, pp. vii + 406.

A bibliography following each chapter in the book furnishes the sources of the original material and provides the instructor with a wealth of supplementary reference.

C. N. COLLOTON.

---

5. "*A Three-foot Shelf of Tests.*"—Thus does the author[1] of this handbook on educational measurements play on Dr. Eliot's phrase and characterize his selection of tests. The book starts out with six brief chapters on elementary statistics. The measurement of mental ability is discussed in seven chapters and the remaining seventeen chapters are devoted to achievement tests in elementary and secondary schools.

---

6. *Discourses on the Psychological and Pedagogical Implications of Adolescence.*[2]—When title, preface and introduction all lead one to expect a practical discussion of "problems" it is indeed disconcerting to be obliged to wade through 356 pages of discourse before encountering the first concrete incident or statement of an actual problem and its solution in practice.

It is equally disturbing to note the wide range of topics discussed in the book and the inadequate treatment of some of the crucial questions. The quotations from G. Stanley Hall, C. H. Johnston, G. H. Palmer and scores of other writers abound, but there is not a footnote in the book and the reader is wearied by the consequent frequency of such indefinite references as the following: "It is well to recall Dr. Hall's summary concerning character . . ." or "The Committee of Ten seems to assume. . . ." Particular works or pages are not mentioned. The frequent use of data from investigations with no mention of the data is as disturbing to the critical reader as the utter lack of such data in the bibliography. For effective use with prospective high school teachers there is an erroneous assumption of educational experience and background, with no urge to mental activity in the form of summaries, suggested problems, illustrative incidents or case studies.

---

[1] Hines, H. C.: "A Guide to Educational Measurements." New York, Houghton Mifflin Company, 1923, pp. 22 + 270.

[2] Pringle, Ralph W.: "Adolescence and High School Problems." New York, D. C. Heath and Company, 1922, pp. x + 386.

7. *A Substitute for the "Reserve Shelf."*—Wide reading in connection with beginners courses in psychology has been facilitated by this compilation and topical organization of some 200 selections from the writings of almost as many outstanding contributors to psychological thought.[1] The compilers have simplified matters for instructors in psychology by supplying thought-provoking problems for use as assignments or exercises in connection with each quotation. If desired the book can be used in connection with a regular textbook. It includes widely differing points of view from which assignments can be made as the purpose, point of view or capacity of the student or instructor may dictate. The organization and content facilitate comparative consideration of various view points and thus stimulate to further inquiry reading and study.

---

8. *A New Book on the Teaching of Reading.*—A book on this subject seems superfluous in view of other recent treatments of the same topic. This notion must be abandoned, however, because of the significance of the organization of this particular publication[2] and the consistent embodiment of a psychological view point in the presentation. Part 1 consists of three introductory chapters. The first is on the aims or objectives of reading instruction. The aims of five teachers, and those given by several authors and educators are cited. The purpose of the chapter is stated. Social and historical data upon which a revision of aim must depend are next included, and specific aims are outlined. There follows a concise conclusion, then a summary of the chapter and a list of readings with specific chapter references. The first paragraph in each successive chapter shows its connection with the preceding chapter, and then continues the discussion according to the scheme of the introductory chapter. Both sides of controversial discussions are presented. Sub-titles are clear forecasts of the content of each section, and abundant concrete examples and instances are used to introduce abstract notions or to exemplify principles. The author shows keen psychological insight of his subject both in his analysis of pupil activities and in the development of his subject.

---

[1] Robinson, Edward S. and Robinson, Florence Richardson: "Readings in General Psychology." Chicago, The University of Chicago Press, 1923, pp. xvi + 674.

[2] Wheat, H. G.: "The Teaching of Reading." New York, Ginn and Company, 1923, pp. xi + 346.

He does not err by serious over- or under-emphasis of topics. About 100 pages are devoted to the reading process and it is in this portion of the book that most of the quotations from scientific studies appear. Part III is entitled "The Course of Study in Reading." A historical chapter is followed by a detailed discussion and practical interpretation of aims and principles by grades. The book concludes with a chapter on remedial work.

---

9. *Impulsive Tendencies Satisfied by Industrial Studies.*—When one considers the divergent and conflicting tendencies of the past 20 years, it is a notable fact in the development of industrial arts education that so few elementary school teachers still attempt to justify industrial studies and experiences on the basis of disciplinary values. These elementary studies of present day industries should share the responsibility with other school subjects for helping pupils to develop appreciative insight and reasoning power in terms of significant interests and actual life needs. These studies are fast replacing the so-called "busy work" or handwork emphasis in industrial courses.

Those who believe that children from six to twelve years of age are concerned mainly with situations and activities similar to those in which adults are engaged, rather than series of formal models or abstract constructions will welcome the new book[1] by Professors Bonser and Mossman. It is a contribution to the literature of industrial arts education. It contains a wealth of suggestive material and makes provision for concrete "outcomes" at the close of each of the several units included for the six elementary grades. For some years the best authors have tried out under varying conditions the principles of selection and organization which they have proposed in Part I of this volume. The results of their experience and observations at the Western Illinois State Normal School, at the Speyer School of Columbia University, and in connection with the work of a large number of cooperating teachers in different public school systems are presented in Part II. This deals with the application of these principles to the activities in which the children are expected to engage.

Throughout the book "constructive work is presented as a means of awakening intellectual inquiries, of giving meanings and values, of cultivating appreciations, and of leading on to further interests on

---

[1] Bonser, F. G. and Mossman, L. C.: "Industrial Arts for Elementary Schools." New York, Macmillan Company, 1923, pp. xi + 491.

levels higher than those apparent at the beginning of the constructions." The introduction to a chapter dealing with the psychology of the industrial arts calls attention to marked instinctive tendencies. Such examples are given as the impulse to manipulative activity, the impulse to investigate, the artistic or æsthetic impulse, and the social impulse. In this broader interpretation of instinct, the authors have had the following in mind: "The impulses of children to investigate and to enjoy, or appreciate, have been utilized quite as much as the impulses to manipulate and construct."

This book can be used advantageously as a handbook by progressive elementary school teachers as well as by those students who are preparing to teach in the elementary grades.

A. H. EDGERTON.

---

10. *Movies in the Schools.*—To the superintendent, principal or teacher who wants to use motion pictures in the school but does not know exactly how to go about it, "Motion Pictures in Education"[1] will prove of real value. Excellent advice is furnished on technical details of projection; and sources of films are indicated. The book will effectively short-circuit the long process of trial and error through which many of us have had to pass in visual instruction.

Modern educational theory would take exception to much of the philosophy of the book. Education is defined in one place as "the imparting and acquisition of knowledge;" in another place as "the harmonious unfolding of all the faculties." Memory is considered as a "faculty" capable of general development. Several of the Pedagogical Principles on page 161 are open to serious criticism. But the suggestions for the use of films in the class-room are much better than the philosophy, and will be of great service to teachers who are struggling to evolve a methodology for the use of this newest of teaching devices.

EDWIN H. REEDER.

---

11. *Teaching Composition in France.*—Phyllis Robbins' translation of Mr. Bezard's intimate account of what he says and does when teaching "Class 2" boys (16 to 18 years of age) to write will particularly please those whose interest in French composition teaching was aroused

[1] Ellis, D. C. and Thornborough, L.: "Motion Pictures in Education." New York, Crowell Company, 1923, pp. xvii + 284.

several years ago by Rollo Walter Brown's *How The French Boy Learns To Write.* In this new book,[1] however, one's interest is likely to center finally in Mr. Bezard, the teacher, rather more than in Mr. Bezard's method. The method, if Mr. Bezard is left out, seems traditional, academic, and what we in America are inclined to call unpsychological. But the teacher in this case has the breadth of view, the sensitiveness to boyish life, the facility of mind, and, above all, the good sense to make any method work. Most of us, at any rate, would like to write as well as do the boys whose themes or parts of themes appear in this volume.

Maybe, too, the American teacher of composition might profitably consider again, in the light of Mr. Bezard and his success, the validity of such ideas as these: The teacher of composition should be a person of outstanding culture and of wide interests. He should be able to write well himself and to demonstrate this for his pupils. Reading experiences may be more valuable for theme material than "personal experiences." The building of character should be a very conscious aim in composition teaching. Classic models are well worth imitation. Themes should be less frequent and more discursive.

M. H. WILLING.

---

12. *A Language Text.*—Book Three of the Cowan, Betz, Charters[2] language series is much more practical in its subject-matter for junior high school use than most of its kind. It stresses functional grammar and ordinary good usage. Tests, reviews, and drill exercises are numerous, but, as must always be the case under book limitations, not nearly numerous enough and not nearly effective enough in design.

M. H. WILLING.

---

[1] Bezard, J.: Translator, Robbins, P.: "My Class in Composition, A Teacher's Diary." Cambridge, Harvard University Press, 1923, pp. xxxi + 268.

[2] Cowan, E. M., Betz A. and Charters, W. W.: "Essential Language Habits." New York, Silver, Burdett and Company, 1923, pp. viii + 439.

DEC 17 1923.

Index and Title Page in This Number

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS:

HAROLD ORDWAY RUGG, *Chairman.*
Lincoln School of Teachers College.
Teachers College, Columbia University.

JAMES CARLETON BELL,
Brooklyn Training School for Teachers.

FRANK NUGENT FREEMAN,
University of Chicago.

ARTHUR IRVING GATES,
Teachers College, Columbia University.

VIVIAN ALLEN CHARLES HENMON,
University of Wisconsin.

RUDOLF PINTNER,
Teachers College, Columbia University.

BEARDSLEY RUML,
Carnegie Corporation, New York City.

LEWIS MADISON TERMAN,
Leland Stanford University.

EDWARD LEE THORNDIKE,
Teachers College, Columbia University.

LAURA ZIRBES, *Assistant Editor.*
Lincoln School, of Teachers College.

VOL. XIV.
No. 9.

DECEMBER, 1923

$4.00 a year.
60c. per copy.

CONTENTS

ON THE IMPROVEMENT IN INTELLIGENCE SCORES FROM FOURTEEN TO EIGHTEEN. *E. L. Thorndike* 513

THE RELATIVE PREDICTIVE VALUES OF CERTAIN INTELLIGENCE AND EDUCATIONAL TESTS TOGETHER WITH A STUDY OF THE EFFECT OF EDUCATIONAL ACHIEVEMENT UPON INTELLIGENCE TEST SCORES. *Arthur I. Gates and Jessie La Salle.* 517

A STUDY OF THE RELATION BETWEEN ABILITY TO LEARN AND INTELLIGENCE AS MEASURED BY TESTS. *O. J. Johnson.* . . . 540

Index and Title Page of Volume XIV (1923). . . . .Facing page 542

SCALES FOR MEASURING JUDGMENT OF ORCHESTRAL MUSIC. *M. R. Trabue* 545

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES 562

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 567

Published Monthly Except June to August by

WARWICK and YORK, Inc.,

York, Pa. *Baltimore, Md.*

ered as Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania, under the Act of March 3, 1879.